Modeling HIV Immune Response and Validation with Clinical Data

H. T. Banks[a,1], M. Davidian[a,2], Shuhua Hu[a,3], Grace M. Kepler[a,4] and E.S. Rosenberg[a,b,5]

[a]Center for Research in Scientific Computation
North Carolina State University
Raleigh, NC 27695-8205
and
[b]Massachusetts General Hospital and Harvard Medical School
I.D. Unit–Gray 5
Boston, MA 02114

March 28, 2007

Abstract

A system of ordinary differential equations is formulated to describe the pathogenesis of HIV infection, wherein certain important features that have been shown important by recent experimental research are incorporated in the model. These include the role of CD4+ memory cells that serve as a major reservoir of latently infected cells, a critical role for T-helper cells in the generation of CD8 memory cells capable of efficient recall response, and stimulation by antigens other than HIV. A stability analysis illustrates the capability of this model in admitting multiple *locally asymptotically stable* (locally a.s.) off-treatment equilibria. The phenomenon of "viral blips" experienced by some patients on therapy with viral load levels suppressed below the detection limit is also investigated. Censored clinical data is used to demonstrate that this model provides reasonable fits to all the patient data available for this study and, moreover, that it exhibits impressive predictive capability.

Key words: HIV, immune response, censored data, inverse problems, viral blips, model prediction, multiple equilibria

[1]email: htbanks@ncsu.edu (corresponding author)
[2]email: davidian@ncsu.edu
[3]email: shu3@ncsu.edu
[4]email: gmkepler@ncsu.edu
[5]email: erosenberg1@partners1.org

1

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **28 MAR 2007**	2. REPORT TYPE	3. DATES COVERED **00-00-2007 to 00-00-2007**

4. TITLE AND SUBTITLE **Modeling HIV Immune Response and Validation with Clinical Data**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **North Carolina State University,Center for Research in Scientific Computation ,Raleigh,NC,27695-8205**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES **37**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

1 Introduction

Since the seminal work of Ho, et al., [18] demonstrated the promise for elucidating HIV disease mechanisms through mathematical modeling, a wide variety of models have been proposed to describe various aspects of in-host HIV infection dynamics (e.g., [1, 2, 3, 12, 16, 25, 26]). The most basic of these models typically include two or three of the key dynamic compartments: virus, uninfected target cells, and infected cells. These compartmental depictions lead to systems of linear or nonlinear ordinary differential equations in terms of state variables representing the concentrations in each compartment and parameters describing viral production and clearance, cell infection and death rate, treatment efficacy, etc. Solutions for the model states yield the time course of viral load and CD4+ counts, for example.

Although such models can be expected only to approximate the myriad processes underlying HIV pathogenesis, when used in conjunction with data as part of designed experiments, these models can be powerful tools in answering questions about the pathogenesis of HIV infection or similar biological processes. Mathematical models can also stimulate further clinical and laboratory research [26]. For example, early applications of linear systems to short-term data on patients undergoing ARV therapy suggested the now widely-held theory of very rapid and constant turnover of viral and infected cell populations [18, 25], contradicting previous assumptions that stable viral and CD4+ concentrations during the clinical latency period of chronic HIV infection are due to absence of significant viral replication.

Some important features of HIV pathogenesis and the cellular immune response that have emerged in recent research include a clearer delineation of the importance of memory $CD4^+$ T-cells as a latent reservoir of HIV [17, 33] and a critical role for T-helper cells in the generation of CD8 memory cells capable of an efficient recall response [4, 8, 19]. The authors in [33] indicate that, even in treated patients who have had no detectable viremia for as long as 7 years, the reservoir decays so slowly that early initiation of Highly Active Anti-Retroviral Therapy (HAART) with the goal of virus eradication is not likely to succeed. This motivates us to develop a model that incorporates these features. In any discussions of mathematical modeling of complex systems it is appropriate to point out that while complex models may be needed to provide accurate descriptions of the underlying dynamics, the models are most useful when they can be compared to clinical and/or experimental data and can also be used for prediction. In developing models for HIV infection and treatment or some other biological phenomenon, this requires a balance between complexity and utility. Hence, in this paper we do not try to formulate a model that reflects all features of cellular immune response as well as all host and viral factors. Instead, we attempt to develop a model that can capture the most salient features of disease progression, that can be supported and validated based on data, one for which parameters can be plausibly estimated, one that has predictive capabilities, and one for which control/drug therapy design is tractable. While the model developed and analyzed here is new, it modifies and extends both conceptually and structurally the predictive model in [3]. That model included both CD4+/viral dynamics as in models discussed in [12] as well as immune response compartments whose importance have been earlier established [10, 24, 39] – see the discussions in [1].

The paper is organized as follows. In Section 2, a system of ordinary differential equations is developed to describe the pathogenesis of HIV and the cellular immune response. In Section 3 we discuss the ability of the model to admit multiple locally a.s. off-treatment equilibria. In Section 4 we investigate a possible mechanism for producing viral blips and elucidate the role of latently-infected CD4+ memory cells and the effect of CD4+ help on CD8+ memory during the ensuing immune response. The expectation maximization algorithm leading to weighted least-squares techniques is employed in Section 5 to fit the model to clinical data with lower limit censoring. The predictive capability of the model is also investigated by using simulation results, with parameters estimated from only half of the longitudinal observations, to predict the immune response in the latter half and comparing these predictions to clinical observations and model results obtained from fitting the full longitudinal data sets. Finally we close with summary conclusions and remarks in Section 6.

2   HIV Model

The model we develop in this paper conceptually modifies and extends the model in [3], wherein two types of target cells (CD4+ T-cells and macrophages), along with their corresponding infected states, free virus, and immune effector cells (CTL) are included in the model. Clinical data fitting results show that the preliminary model of [3] provides reasonable fits to most patient data and has impressive predictive capability when comparing model simulations, with parameters based on estimation using only half of the longitudinal observations, to the full longitudinal data sets. However, that model does not incorporate some important features of HIV pathogenesis and the cellular immune response, such as CD4+ memory cells as the major reservoir of latently infected cells and a critical role for T-helper cells in the generation of CD8 memory cells capable of an efficient recall response. To incorporate these important features, we thus seek a model that includes some measure of CD4+ T-helper cells, infected memory CD4+ T-cells and HIV-specific memory CD8+ T-cells. To retain the simplicity of the model, secondary target cells, such as macrophages, are not included as a compartment in our new model. It is worth noting that omitting the secondary target cells should not affect our clinical data fitting and predictive capabilities since this type of cell, even though it is very important at the beginning of infection, does not contribute significantly to the virus pool in the long run. The model compartments are illustrated in Table 1, wherein the resting CD4+ T-cells (T_2) are assumed to include naive CD4+ T-cells and memory CD4+ T-cells. This is reasonable since these two types of cells have similar behavior such as longer life spans and distribution in the lymphoid tissue. Once these resting CD4+ T-cells become activated, either through antigen priming of naive cells or reactivation of memory cells, they are more susceptible to HIV infection than resting cells and suffer elevated mortality. Hence, we include these activated naive cells and reactivated memory cells in the other compartment as activated CD4+ T-cells (T_1). Infected resting and activated cells are represented by the T_2^* and T_1^* states, respectively. A schematic of this new model is depicted in Fig. 1.

The corresponding compartmental ordinary differential equation (ODE) model for in-host

3

states	unit	description
T_1	cells/μl-blood	uninfected activated CD4+ T-cells
T_1^*	cells/μl-blood	infected activated CD4+ T-cells
T_2	cells/μl-blood	uninfected resting CD4+ T-cells
T_2^*	cells/μl-blood	infected resting CD4+ T-cells
V_I	RNA copies/ml-plasma	infectious free virus
V_{NI}	RNA copies/ml-plasma	non-infectious free virus
E_1	cells/μl-blood	HIV-specific effector CD8+ T-cells
E_2	cells/μl-blood	HIV-specific memory CD8+ T-cells

Table 1: Model States.
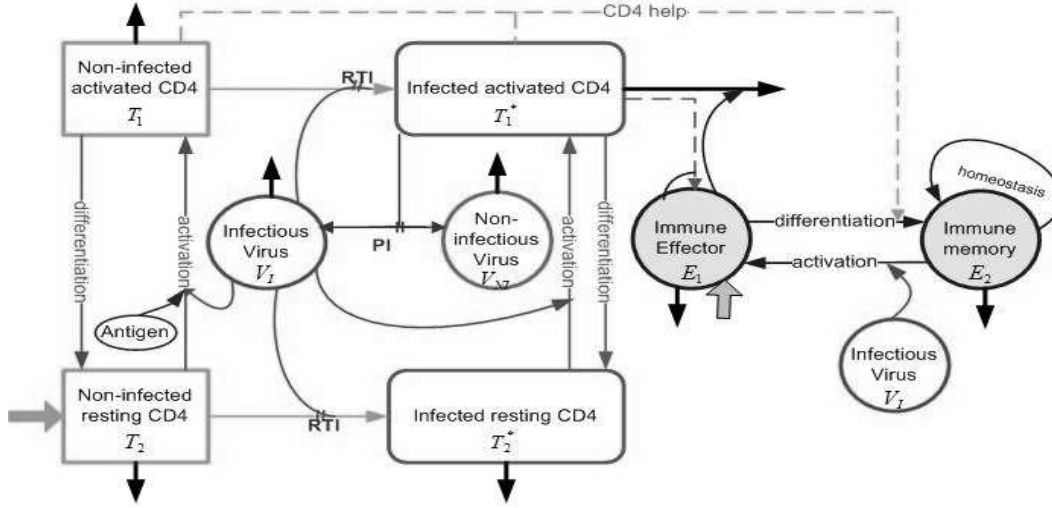
Figure 1: Flow chart of model (2.1), where PI and RTI denote protease inhibitor and reverse transcriptase inhibitor, respectively.

HIV infection dynamics is based on balance laws and is given by

$$\dot{T}_1 = -d_1 T_1 - (1 - \xi_1(t))k_1 V_I T_1 - \gamma_T T_1 + p_T \left(\frac{a_T V_I}{V_I + K_V} + a_A \right) T_2,$$

$$\dot{T}_1^* = (1 - \xi_1(t))k_1 V_I T_1 - \delta T_1^* - m E_1 T_1^* - \gamma_T T_1^* + p_T \left(\frac{a_T V_I}{V_I + K_V} + a_A \right) T_2^*,$$

$$\dot{T}_2 = \lambda_T \frac{K_s}{V_I + K_s} + \gamma_T T_1 - d_2 T_2 - (1 - f\xi_1(t))k_2 V_I T_2 - \left(\frac{a_T V_I}{V_I + K_V} + a_A \right) T_2,$$

$$\dot{T}_2^* = \gamma_T T_1^* + (1 - f\xi_1(t))k_2 V_I T_2 - d_2 T_2^* - \left(\frac{a_T V_I}{V_I + K_V} + a_A \right) T_2^*, \tag{2.1}$$

$$\dot{V}_I = (1 - \xi_2(t))10^3 N_T \delta T_1^* - c V_I - 10^3[(1 - \xi_1(t))\rho_1 k_1 T_1 + (1 - f\xi_1(t))\rho_2 k_2 T_2]V_I,$$

$$\dot{V}_{NI} = \xi_2(t)10^3 N_T \delta T_1^* - c V_{NI},$$

$$\dot{E}_1 = \lambda_E + \frac{b_{E1} T_1^*}{T_1^* + K_{b1}} E_1 - \frac{d_E T_1^*}{T_1^* + K_d} E_1 - \delta_{E1} E_1 - \gamma_E \frac{T_1 + T_1^*}{T_1 + T_1^* + K_\gamma} E_1 + \frac{p_E a_E V_I}{V_I + K_V} E_2,$$

$$\dot{E}_2 = \gamma_E \frac{T_1 + T_1^*}{T_1 + T_1^* + K_\gamma} E_1 + \frac{b_{E2} K_{b2}}{E_2 + K_{b2}} E_2 - \delta_{E2} E_2 - \frac{a_E V_I}{V_I + K_V} E_2,$$

with an initial condition vector

$$[T_1(0), T_1^*(0), T_2(0), T_2^*(0), V_I(0), V_{NI}(0), E_1(0), E_2(0)]^T.$$

Here the factors 10^3 are introduced to convert between microliter and milliliter scales, preserving the units from some of the earlier published papers [1, 10]. The treatment factors $\xi_1(t) = \epsilon_1 u(t)$ and $\xi_2(t) = \epsilon_2 u(t)$ represent the effective treatment impact, consisting of efficacy factors ϵ_1 modeling the relative effectiveness of reverse transcriptase inhibitor (RTI), ϵ_2 describing the relative effectiveness of protease inhibitor (PI), and a time-dependent treatment function $u(t)$ ($0 \leq u(t) \leq 1$) representing HAART drug level, where $u(t) = 0$ is fully off and $u(t) = 1$ is fully on. Since HIV treatment is nearly always administered as combination therapy, we do not consider the possibility of monotherapy, even for a limited period of time, though this could be implemented by considering separate treatment functions.

The input term $\lambda_T \frac{K_s}{V_I + K_s}$ for the T_2 compartment is used to account for the source rate of uninfected resting CD4+ T-cells. This term depends on the viral load level since the thymus production can be diminished if the viral load is too high [22]. To limit the introduction of additional parameters, we assume that uninfected and infected resting CD4+ T-cells (T_2 and T_2^*, respectively) have the same natural death rate d_2. We remark that activated CD4+ T-cells have a higher natural death rate than resting memory and naive cells, and we use d_1 to denote the natural death rate of uninfected activated CD4+ T-cells T_1. The immune effector cells E_1 remove infected activated cells CD4+ T-cells T_1^* from the system by cell lysis with a rate m. However, immune effector cells do not remove infected resting cells T_2^*, since these cells are in a quiescent state where the virus is not replicating and, thereby, escape the detection of the immune effector cells. These infected resting cells are assumed to become targets for lysis only after activation [9].

The infected activated cells T_1^* result from encounters between uninfected activated cells T_1 and free infectious virus V_I with infection rate k_1. The resulting term $k_1 V_I T_1$ is modified by a factor $1 - \xi_1(t)$ to account for RTI treatment. Infection of the resting T-cell compartment T_2,

5

which is comprised of both memory and naive CD4+ T-cells, can occur in a number of ways. First, the most commonly transmitted R5 virus form of HIV-1 that utilize the chemokine receptor CCR5 can enter a subset of resting memory cells that express sufficient levels of CCR5 to support infection [9]. In addition, the X4 form of the virus can infect resting CD4+ T-cells, whether they belong to the naive or memory subsets. However, infection of naive and memory cells through these routes occurs much less frequently than infection of T_1, and, once infected, these cells often do not progress to a long-term stably-infected state in which the virus is integrated into the host DNA. In addition, it has been shown that infected naive CD4+ T-cells do not significantly contribute to the pool of infected resting CD4+ T-cells [11]. Hence, the term $(1 - f\xi_1(t))k_2 V_I T_2$ is used to represent the infection process that results from encounters between the uninfected resting CD4+ T-cells and free virus V_I, but with an infection rate $k_2 < k_1$ to account for a significantly lower rate of infection as compared to activated CD4+ T-cells. The treatment factor $\xi_1(t)$ is potentially more effective in T_1 than in T_2, where the efficacy is modelled by $f\xi_1(t)$ with $0 \le f \le 1$.

A much more stable form of latent infection arises when activated CD4+ T-cells that have integrated HIV-1 DNA survive long enough to revert back to resting memory state, and latently infected resting CD4+ T-cells with integrated HIV-1 DNA are present in all infected individuals but only at low frequency [9, 14]. Hence, the terms involving $\gamma_T T_1^*$ are included in the model to account for the phenomenon of differentiation of infected activated CD4+ T-cells into infected memory or resting CD4+ T-cells T_2^* at rate γ_T. For simplicity, the rate at which uninfected activated CD4+ T-cells T_1 differentiate into uninfected resting CD4+ T-cells T_2 is also assumed to be γ_T; the model could be extended easily to the case with different differentiation rates.

As the authors of [9]concluded, there is turnover in the latent reservoir when patients are viremic, but the degree of turnover depends on the level of viremia. We thus assume that the activation of infected HIV-specific resting CD4+ T-cells T_2^* depends on the virus concentration with a half-saturation constant K_V. Hence, the terms involving $\frac{a_T V_I}{V_I + K_V} T_2^*$ are used to represent the activation of infected HIV-specific resting CD4+ T-cells with maximum activation rate of a_T. Again to preserve the simplicity of this model, we assume that activation of uninfected HIV-specific resting CD4+ T-cells T_2 also depends on the virus concentration, with a half saturation constant K_V, and that the maximum activation rate is also a_T. Thus, the terms involving $\frac{a_T V_I}{V_I + K_V} T_2$ represent the activation of uninfected HIV-specific resting CD4+ T-cells. In order to incorporate the activation of resting CD4+ T-cells by some non-HIV antigen and preserve the simplicity of the model, we include the simple terms $a_A T_2$ and $a_A T_2^*$ into our model to describe this phenomenon, with a_A being the activation rate by non-HIV antigen. The parameter a_A here can be utilized as a constant to represent a chronic level of infection or as a function of time t to describe infections that are cleared by the body. These activation terms represent losses to the T_2 and T_2^* compartments, with corresponding gain terms for the T_1 and T_1^* compartments. However, the gain terms for T_1 and T_1^* include a multiplicative factor p_T to account for the net proliferation due to clonal expansion and programmed contraction. For simplicity, we assume that uninfected and infected CD4+ T-cells have the same expansion factor p_T; again this can be readily extended to include processes with different expansion factors.

Virus in the reservoir T_2^* of infected resting CD4+ T-cells is latent and no virus can be produced by these cells unless they are activated [9]. Hence, free virus particles V_I are produced only by activated infected CD4+ T-cells during viral budding leading up to viral produced lysis δT_1^* of the CD4+ T-cells. The parameter N_T accounts for the number of RNA copies produced during this process in the viral source term $(1 - \xi_2(t))10^3 N_T \delta T_1^*$. In addition to a natural clearance rate c, we also include terms $10^3[(1 - \xi_1(t))\rho_1 k_1 T_1$ and $(1 - f\xi_1(t))\rho_2 k_2 T_2]V_I$ in the free virus compartment V_I to account for the removal of free virus that takes place when free virus infects T_1 and T_2. We make the simplifying assumption that $\rho_i = 1 \dfrac{\text{copies ml-blood}}{\text{cells ml-plasma}}$, i.e., one free virus particle is responsible for each new infection. This could be adapted easily for multiple virus particles being responsible for each new infection by choosing $\rho_i > 1$. Since clinical measurements of viral load do not differentiate between infectious and non-infectious virus, we include a compartment in the model for tracking the amount of non-infectious virus V_{NI}. The action of a protease inhibitor, resulting in the production of non-infectious virus V_{NI} by infected cells is modeled by ξ_2. It should be noted that the inclusion of this additional state does not affect the dynamics of the other state variables.

The source term λ_E, the constant death term δ_{E1}, and the nonlinear infected cell-dependent birth $\frac{b_{E1}T_1^*}{T_1^*+K_{b1}}E_1$ and death $\frac{d_E T_1^*}{T_1^*+K_d}E_1$ terms in the E_1 compartment are adopted from the model in [2, 3], where the authors suggested that, by including such terms in the immune effector compartment, the model can admit multiple stable off-treatment steady states and exhibit transfer between "healthy" and "unhealthy" locally stable steady states via optimal or suboptimal structure treatment interruptions (STI) therapies. This makes it a good candidate for our investigation. Memory CD8+ T-cells are also subject to strict homeostatic control [34]; background expansion of memory cells through intermittent cell division being countered by an equivalent level of cell death. Hence, we include the term $\frac{b_{E2}K_{b2}}{E_2+K_{b2}}E_2 - \delta_{E2}E_2$ for homeostatic regulation in the E_2 compartment, similar to that used in [37]. In the homeostatic regulation, b_{E2} represents the maximum proliferation rate and δ_{E2} corresponds to the death rate, where the proliferation signal decreases linearly with population size.

The term $\gamma_E \frac{T_1+T_1^*}{T_1+T_1^*+K_\gamma}E_1$ in the model is used to include the essential role that activated CD4+ T-cells play in the generation of memory CD8+ T-cells during the priming phase, where parameter K_γ is a half-saturation constant and γ_E is the maximum rate at which E_1 differentiates into E_2. Since depletion of CD4+ cells has a minimal effect during the recall response [32, 38], the term $\frac{a_E V_I}{V_I+K_V}E_2$ for reactivation of memory CD8+ T-cells is independent of CD4+ T-cell help. Similar to the activation of HIV-specific resting CD4+ T-cells, we assume that activation of HIV-specific memory CD8+ T-cells also depends on the virus concentration. For simplicity, we use the same half-saturation constant K_V for the activation of memory CD8+ T-cells. Since CD8+ T-cells tend to divide sooner and to have a faster rate of cell division than CD4+ T-cells [31], we use a different parameter p_E to account for the net proliferation due to clonal expansion and programmed contraction of activated CD8+ T-cells in the E_1 compartment.

3 Off-treatment Stability Analysis

Our model choice is partly motivated by its admission of multiple locally a.s. off-treatment equilibria (corresponding to $\epsilon_1 = 0$ and $\epsilon_2 = 0$), which is important in the sense that we seek to investigate the possibility of improvement of long-term health by using STI during acute infection to affect a change from an "unhealthy" equilibrium to a "healthy" one. A general analysis of the equilibria of the model (2.1) and their stability properties is challenging due to the complexity of the system and is generally not solvable in analytic form. Hence, in this section we illustrate the existence of multiple locally a.s. equilibria through specific examples and numerically investigate the behavior of these equilibria with respect to changes in parameters and initial conditions.

Uninfected equilibrium in a healthy person ($a_A = 0$). In the absence of non-HIV antigen ($a_A = 0$), it is easy to see that model (2.1) has the following off-treatment uninfected equilibrium:

$$(0, 0, \lambda_T/d_2, 0, 0, 0, \lambda_E/\delta_{E1}, 0). \tag{3.1}$$

Substituting the above equilibrium into the Jacobian matrix of model (2.1), we find that the eigenvalues of this matrix are given by $-(\delta\delta_{E1} + m\lambda_E + \gamma_T\delta_{E1})/\delta_{E1}$, $-c$, $-d_2$, $-d_2$, $-(cd_2 + 1000k_2\lambda_T)/d_2$, $-\delta_{E1}$, $b_{E2} - \delta_{E2}$ and $-d_1 - \gamma_T$. Hence, if $b_{E2} < \delta_{E2}$ then all the eigenvalues of this Jacobian matrix are negative, which implies that under this case equilibrium (3.1) is locally a.s. On the other hand, if $b_{E2} > \delta_{E2}$, then we obtain a different uninfected equilibrium:

$$(0, 0, \lambda_T/d_2, 0, 0, 0, \lambda_E/\delta_{E1}, (b_{E2}/\delta_{E2} - 1)K_{b2}). \tag{3.2}$$

In this case, we find that the eigenvalues of Jacobian matrix are given by $-(\delta\delta_{E1} + m\lambda_E + \gamma_T\delta_{E1})/\delta_{E1}$, $-c$, $-d_1 - \gamma_T$, $-(cd_2 + 1000k_2\lambda_T)/d_2$, $-\delta_{E1}$, $-\delta_{E2}(b_{E2} - \delta_{E2})/b_{E2}$, $-d_2$ and $-d_2$. Hence, equilibrium (3.2) is also locally a.s. Therefore, for the case $a_A = 0$, model (2.1) always has a locally a.s. off-treatment uninfected equilibrium. The existence of a locally a.s. uninfected equilibrium is biologically reasonable in light of research documenting the existence of some sex workers who, though regularly exposed to HIV-contaminated body fluids, remain HIV-negative [28, 29]. We note that this feature of our new model is not present in the earlier model investigated in [3].

Multiple equilibria in a healthy person ($a_A = 0$). We next illustrate the existence of multiple equilibria with specific examples when $a_A = 0$ and the values of all the other parameters as specified in Table 2, unless otherwise stated. Since values for most of the parameters in this model can not be found in the literature, the values listed in Table 2 are chosen for model illustration purposes. When determining equilibria for this complex model there are usually many unstable or physically-meaningless (eg., negative state values) equilibria. In the discussion that follows, unless otherwise stated, we focus our attention on the locally a.s., physically-meaningful equilibria only.

In particular, we wish to explore the effect that CD4+ memory cell activation plays on the equilibrium states available to the system. Since, for economy, we have included both naive and memory cells in a single resting CD4+ T-cell compartment T_2 (or T_2^* for the

parameter	value	parameter	value	parameter	value
d_1	$0.02 \dfrac{1}{\text{day}}$	ϵ_1	$\in [0,1]$	k_1	$10^{-5} \dfrac{\text{ml-plasma}}{\text{copies} \cdot \text{day}}$
δ	$0.7 \dfrac{1}{\text{day}}$	m	$0.01 \dfrac{\mu\text{l-blood}}{\text{cells} \cdot \text{day}}$	λ_T	$7 \dfrac{\text{cells}}{\mu\text{l-blood} \cdot \text{day}}$
d_2	$0.005 \dfrac{1}{\text{day}}$	K_V	$100 \dfrac{\text{copies}}{\text{ml-plasma}}$	K_s	$10^5 \dfrac{\text{copies}}{\text{ml-plasma}}$
f	0.34	γ_T	$0.005 \dfrac{1}{\text{day}}$	N_T	$100 \dfrac{\text{copies} \cdot \text{ml-blood}}{\text{cells} \cdot \text{ml-plasma}}$
ϵ_2	$\in [0,1]$	c	$13 \dfrac{1}{\text{day}}$	λ_E	$0.001 \dfrac{\text{cells}}{\mu\text{l-blood} \cdot \text{day}}$
b_{E1}	$0.3 \dfrac{1}{\text{day}}$	d_E	$0.25 \dfrac{1}{\text{day}}$	K_{b1}	$0.1 \dfrac{\text{cells}}{\mu\text{l-blood}}$
a_A	$[0,1] \dfrac{1}{\text{day}}$	δ_{E1}	$0.1 \dfrac{1}{\text{day}}$	K_d	$0.5 \dfrac{\text{cells}}{\mu\text{l-blood}}$
a_T	$0.008 \dfrac{1}{\text{day}}$	p_T	1.2	k_2	$10^{-9} \dfrac{\text{ml-plasma}}{\text{copies} \cdot \text{day}}$
a_E	$0.1 \dfrac{1}{\text{day}}$	p_E	3	K_γ	$10 \dfrac{\text{cells}}{\mu\text{l-blood}}$
γ_E	$0.01 \dfrac{1}{\text{day}}$	b_{E2}	$0.001 \dfrac{1}{\text{day}}$	K_{b2}	$100 \dfrac{\text{cells}}{\mu\text{l-blood}}$
δ_{E2}	$0.005 \dfrac{1}{\text{day}}$				

Table 2: Parameter values used in model (2.1) to illustrate the existence of multiple equilibria. Note that for this parameter set $b_{E2} < \delta_{E2}$.

			a_T				
			8×10^{-3}		8×10^{-2}		1×10^{-3}
state	EQ_1	EQ_2	EQ_3	EQ_2	EQ_3	EQ_2
T_1	0	184.3	13.13	352.1	13.12	40.27
T_1^*	0	0.05621	5.330	0.05582	7.659	0.05935
T_2	1400	731.2	424.4	142.8	59.36	1247
T_2^*	0	0.04393	2.984	0.005110	0.4801	0.1187
V_I	0	265.1	28410	236.5	40830	310.0
V_{NI}	0	0	0	0	0	0
E_1	0.01	799.2	0.04574	1422	0.03872	140.0
E_2	0	98.32	0.002863	184.5	0.002520	14.07

Table 3: Off-treatment steady states for model (2.1) with $a_A = 0$ day^{-1}, $a_T = 8 \times 10^{-3}$, 8×10^{-2}, and 1×10^{-3} day^{-1}, and the values of all the other parameters as specified in Table 2. Non-physical and unstable steady states are omitted.

infected case), the activation factor $\frac{a_T V_I}{V_I + K_V}$ represents both activation of naive cells and the reactivation of memory cells. In the limit of large V_I the maximum activation rate is a_T. In the treatment that follows, we examine the role of a_T on the equilibrium states available to the system. We first examine the off-treatment equilibria (listed in Table 3) for the case $a_T = 8 \times 10^{-3}$ day^{-1}. We can see that in this and each of the other two cases we consider, there is an uninfected equilibrium similar to the one in (3.1), which we designate as EQ_1. In addition to the uninfected equilibrium, we have two infected equilibria, designated EQ_2 and EQ_3, where EQ_2 represents a "healthy" steady state with immune control of the viral infection, restoration of CD4+ T-cell counts (915.6 cells/μl-blood), and a strong CD8+ immune response (897.5 cells/μl-blood). Equilibrium EQ_3 represents an "unhealthy" steady state corresponding to a dangerously high viral load set point, lower CD4+ T-cell counts (445.8 cells/μl-blood), and a much lower CD8+ immune response (0.04860 cells/μl-blood).

If we begin simulations with the initial conditions $(0, 0, 1400, 0, \nu, 0, 0.01, 0)$, i.e., a nontrival amount ν of infectious free virus, then the solution of model (2.1) converges to either the uninfected equilibrium EQ_1 or the "unhealthy" infected equilibrium EQ_3 (see Fig. 2). The particular steady state that the model converges to depends on the viral load level ν. Simulation results reveal that the solution will not converge to EQ_3 until the value of ν is close to 6576 copies/ml-plasma.
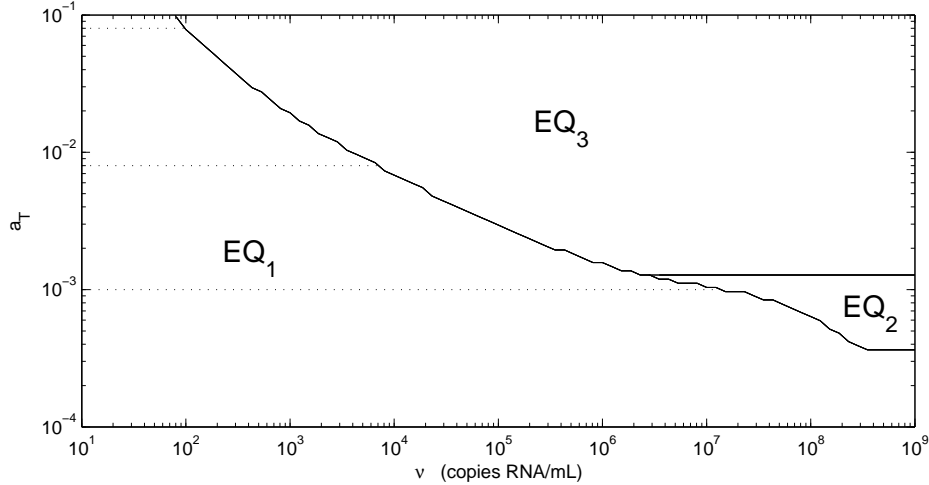
Figure 2: Phase diagram showing equilibrium attained as a function of the initial viral load ν and the parameter a_T. It should be emphasized that this plot is only applicable for the particular initial condition $(0, 0, 1400, 0, \nu, 0, 0.01, 0)$.

If we set the value of a_T larger, such as 8×10^{-2} day^{-1}, and keep the values of all the other parameters the same, then we still have three locally a.s. off-treatment equilibria: uninfected, "healthy" infected and "unhealthy" infected, where the uninfected equilibrium is the same as before (Table 3). However, the "unhealthy" infected equilibrium (EQ_3) in this case has a much higher viral load, a much lower CD4+ T-cell count (80.62 cells/μl-blood), and a lower immune response (0.04124 cells/μl-blood) than the EQ_3 when $a_T = 8 \times 10^{-3}$ day^{-1}. Even though the "healthy" infected equilibrium EQ_2 has a slightly lower viral load level and a much higher CD8+ immune response (1607 cells/μl-blood) than the EQ_2 when

$a_T = 8 \times 10^{-3}$ day^{-1}, it has degraded with respect to the CD4+ T-cell count (495.0 vs. 915.6 cells/μl-blood).

If we now start simulations with $a_T = 8 \times 10^{-2}$ day^{-1} and the initial conditions $(0, 0, 1400, 0, \nu, 0, 0.01, 0)$, then the solution of model (2.1) also converges to either uninfected equilibrium or "unhealthy" infected equilibrium depending on ν. But in this case, the solution converges to its corresponding "unhealthy" infected equilibrium with ν around 100.3 copies/ml-plasma (see Fig. 2). Hence, the viral load necessary for infection depends on the maximum activation rate a_T; low values of a_T require a larger viral load to reach an infected state. Furthermore, when infected, smaller values of a_T produce "healthier" infections than infections with larger a_T values, in terms of CD4+ T-cell counts and viral load level.

If we set the maximum activation rate of resting CD4+ T-cells lower, such as $a_T = 1 \times 10^{-3}$ day^{-1}, and keep the values of all the other parameters to be the same as those illustrated in Table 2, then we find that we have only two locally a.s. off-treatment equilibria: the uninfected steady state (EQ$_1$) and "healthy" infected steady state (EQ$_2$), where again the uninfected steady state is the same as with $a_T = 8 \times 10^{-3}$ day^{-1} (Table 3). Even though the "healthy" infected steady state has a higher viral load level (310 copies/ml-plasma) and lower immune response (154 cells/μl-blood) than it does when $a_T = 8 \times 10^{-3}$ day^{-1}, the immune response still controls the viral load to maintain it below a detection limit of 400 copies/ml-plasma. This equilibrium also has much higher CD4+ T-cell counts (1290 cells/μl-blood) than the case when $a_T = 8 \times 10^{-3}$ day^{-1}. If we start the simulation with the initial conditions $(0, 0, 1400, 0, \nu, 0, 0.01, 0)$, then the solution of model (2.1) converges either to its uninfected equilibrium or "healthy" infected equilibrium based on the value of ν, and it converges to this "healthy" infected equilibrium with ν around 1.527×10^7 copies/ml-plasma (Fig. 2).

The examples above demonstrate the existence of multiple off-treatment equilibria and illustrate that changing the value of a parameter, such as the maximum activation rate a_T, has an effect on both the number and "health" characterization of the equilibria. In addition, the initial conditions of the system determine which equilibrium is attained after initial infection. Across a population, the parameter values can be expected to vary to represent different host factors and host-virus interaction rates. In the analysis of parameter a_T, we find that a person with a lower a_T value requires a larger viral load in order to get infected and, once infected, attains a "healthier" set point outcome, in terms of their CD4+ T-cell counts and viral load level, than those with higher values of a_T. Analysis with other parameters, such as the expansion factor p_T, reveals similar behavior.

Multiple equilibria in an unhealthy person $(a_A \neq 0)$**.** In order to investigate whether the activation due to a non-HIV antigen can affect the number of physical equilibria and their local asymptotic stability, we performed simulation with $a_A = 10^{-5}$ day^{-1} and the values of all the other parameters as specified in Table 2. All the locally a.s., physical, off-treatment equilibria are tabulated in Table 4, which indicates that we still have three locally a.s. equilibria: an uninfected steady state EQ_1, a "healthy" infected steady state EQ_2 and an "unhealthy" infected steady state EQ_3. We also observe that EQ_1 has a nonzero E_2 even

though $b_{E2} < \delta_{E2}$ and a nonzero T_1, which is not observed for equilibrium EQ_1 when $a_A = 0$ (Table 3). Otherwise, the equilibria for the two cases ($a_T = 8 \times 10^{-3}$ day^{-1} and $a_A = 0$ or 1×10^{-5} day^{-1}) are very similar.

| | a_A | | | | | |
| | 1×10^{-5} | | | 1×10^{-2} | | |
state	EQ_1	EQ_2	EQ_3	EQ_1	EQ_2	EQ_3
T_1	0.6710	184.5	13.13	266.7	277.2	13.12
T_1^*	0	0.05621	5.332	0	0.05592	6.567
T_2	1398	730.6	424.0	555.6	404.2	228.2
T_2^*	0	0.04388	2.982	0	0.01835	1.777
V_I	0	265.0	28420	0	248.2	35010
V_{NI}	0	0	0	0	0	0
E_1	0.009938	799.9	0.04573	0.009121	1160	0.04113
E_2	0.001562	98.41	0.002862	0.02198	147.6	0.002630
local stability	a.s.	a.s.	a.s.	unstable	a.s.	a.s.

Table 4: Off-treatment steady states for model (2.1) with $a_A = 1 \times 10^{-5}$ and 1×10^{-2} day^{-1} and the values of all the other parameters as specified in Table 2. Non-physical and unstable steady states are omitted, except in the case of EQ$_1$ when $a_A = 1 \times 10^{-2}$ day^{-1}, which is unstable. Note that $a_T = 8 \times 10^{-3}$ day^{-1} for this analysis.

If we start simulations with the initial conditions $(0.6710, 0, 1398, 0, \nu, 0, 0.009938, 0.001562)$, then the solution of model (2.1) converges either to its uninfected equilibrium EQ_1 or "unhealthy" infected equilibrium EQ_3 based on the value of ν, and it will not converge to EQ_3 until ν is close to 4000 (Fig. 3), as compared to 6576 copies/ml-plasma when $a_A = 0$. Hence, a person infected with a non-HIV virus becomes HIV-infected at lower viral loads than when no other infection is present.

If we take the value of a_A larger, such as $a_A = 1 \times 10^{-2}$ day^{-1}, then the uninfected equilibrium is no longer stable, but we still have the locally a.s. "healthy" and "unhealthy" steady states designated as EQ_2 and EQ_3 in Table 4, respectively. If we start simulations with the initial conditions $(266.7, 0, 555.6, 0, \nu, 0, 0.009121, 0.02198)$, then the solution of model (2.1) converges to EQ_3, regardless of the initial viral load ν (Fig. 3). This means that when the system is in the state EQ_1, introduction of even the smallest amount of virus will cause the system to converge to the "unhealthy" equilibrium EQ_3.

Furthermore, Table 4 indicates that as the activation rate by non-HIV antigen a_A becomes larger, the locally a.s. "healthy" steady state has degraded, in terms of the total CD4+ T-cells counts (915.1 vs. 681.1 cells/μl-blood), while the viral load has improved (265.0 vs. 248.2 copies/ml-plasma), as well as the immune response (898.4 vs. 1307 cells/μl-blood). In addition, the locally stable "unhealthy" steady state has worsened, in terms of a much lower CD4+ T-cell count (445.4 vs. 249.4 cells/μl-blood), much higher viral load set point (28420
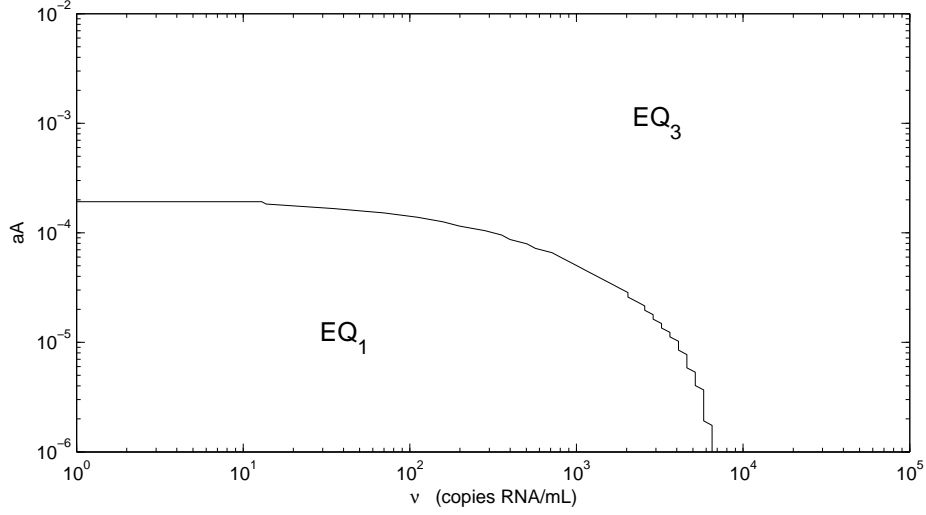
Figure 3: Phase diagram showing equilibrium attained as a function of the initial viral load ν and the parameter a_A. Initial conditions for simulations, corresponding to the uninfected equilibrium EQ_1, are calculated for each value of a_A. Note that $a_T = 8 \times 10^{-3}$ day^{-1} for these simulations.

vs. 35010 copies/ml-blood), and much lower immune response (0.04859 vs. 0.04376 cells/μl-blood). Therefore, at this point increasing a_A leads to similar behavior as does increasing a_T when $a_A = 0$.

The examples above summarize an investigation as to the effect of a non-HIV infection ($a_A \neq 0$) on the outcome of an HIV-infection. We find that an unhealthy person ($a_A \neq 0$) requires a smaller initial viral load ν to sustain an HIV infection and, once infected, attains a worse outcome, in terms of CD4+ T-cell counts and viral load levels, than a healthy person ($a_A = 0$). We find that multiple equilibria, including an uninfected equilibrium, exist in the presence of a small non-HIV infection ($a_A = 1 \times 10^{-5}$ day^{-1}). However, for larger levels of non-HIV infection ($a_A = 1 \times 10^{-2}$ day^{-1}), the uninfected equilibrium is unstable and even the smallest initial viral load will lead to an "unhealthy" equilibrium state (Fig. 3).

4   Viral Blips

Adherence to a regimen of HAART suppresses the viral loads of most infected HIV patients below the level of detection (<400 or <50 copies/ml-plasma depending on the assay used) by standard assays . However, a number of these well-suppressed patients experience unexplained "viral blips" while on therapy [15, 20]. In a study [15] of 123 patients, these viral blips are estimated to have a duration of approximately two to three weeks with a mean blip amplitude of 158 ± 132 copies/ml-plasma. Furthermore, it is observed that the blip frequency inversely correlates with CD4$^+$ T-cell counts. In this section, we investigate how infection with a non-HIV antigen can lead to viral blips for those patients who are on treatment and

13

have successfully suppressed their viral loads to undetectable levels. To do so, we use the states of a locally a.s. "healthy" infected equilibrium $(201, 0.056, 730, 0.040, 288, 0, 240, 29.2)$, corresponding to $\epsilon_1 = 0.7$, $\epsilon_2 = 0$, $a_A = 0$, and the values of all the other parameters as specified in Table 2, as the initial conditions for simulations in which we allow parameter a_A to be a function of time t, while keeping all other parameter values to be the same. As we drive the system (2.1) with the non-HIV infection $a_A(t)$ we monitor the states of the system, particularly the viral load. While the cause of viral blips has not yet been resolved, one proposed mechanism [20] posits that viral blips could be due to an increase in activated CD4+ cells as a result of secondary infection or vaccination. Our approach, introducing a time-dependent non-HIV infection $a_A(t)$ is consistent with this proposed mechanism and demonstrates that our model supports such a scenario.

Viral blip caused by a single non-HIV infection. Figure 4(a) illustrates the case of a one-time non-HIV infection occurring on days 20 through 50, with a peak value of 0.003 day^{-1} (days 30 to 40). Figure 4(b) depicts the HIV viral progression $V_I(t)$ resulting from this infection. In Fig. 4(b) we see that there is a small delay before the viral load begins to increase. This delay is reasonable since there is no activation term (a_A) in the V_I compartment of model (2.1) and the viral load can only increase after T_1^* increases. About 10 days after the start of the infection (day 30), the viral load rises above the detection limit (400 copies/ml-plasma), reaching its peak value 15 days post-infection (day 35). This viral blip drops below the detection limit 20 days post-infection (day 40), before the infection has completely cleared.

(a)

(b)
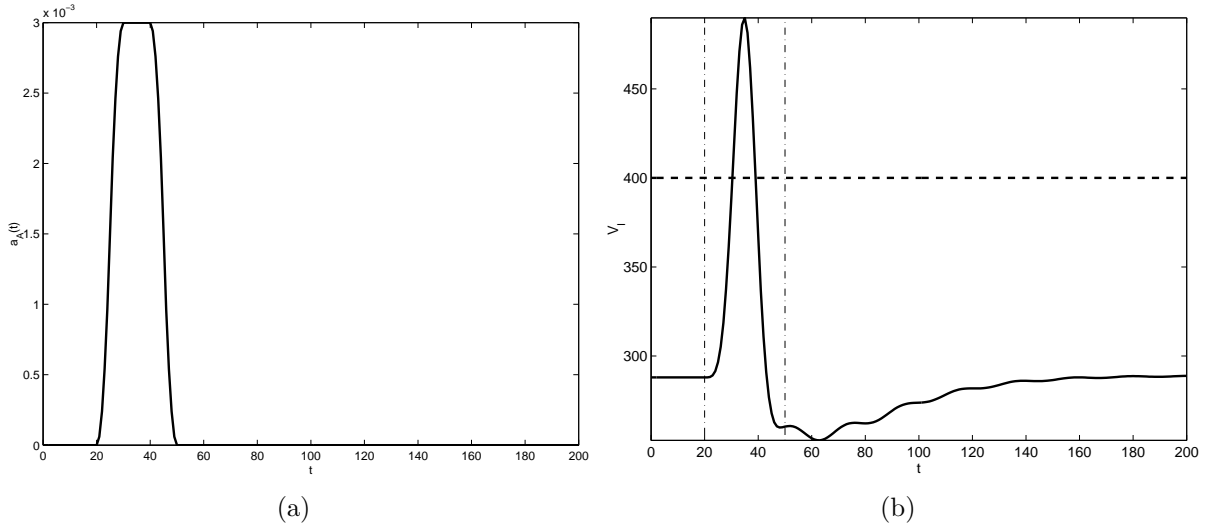
Figure 4: (a) Activation rate $a_A(t)$ by a non-HIV antigen. The duration of the infection is 30 days; (b) Viral load $V_I(t)$ (solid line), censored data level (horizontal dashed line), and infection start and stop times (vertical dash-dot lines).

To further investigate the dynamics of the viral blip, we plot the other model compartments for the example shown in Fig. 4 (we omit compartment V_{NI}, since, with $\epsilon_2 = 0$ and $V_{NI}(0) = 0$, $V_{NI} \equiv 0$). In Fig. 5, it can be seen that the T-cell compartments (T_1, T_1^*, T_2, and T_2^*)

14

respond at the start of the infection (day 20). This is not surprising because the non-HIV antigen activation term a_A appears in the dynamical equations for these compartments in model (2.1). It is also reasonable that the resting (naive and memory) T-cell compartments (T_2 and T_2^*) initially decrease, while the activated T-cell compartments (T_1 and T_1^*) increase, since, at the start of the infection, resting cells are reactivated (memory) to fight the infection or activated through infection (naive). There is also a short delay between the peak value of T_2^* and the peak value of T_1^* which is likely due to the time delay in the differentiation of T_1^* into T_2^*. Comparing compartment T_1^* of Fig. 5 with compartment V_I in Fig. 4, we find that these two compartments exhibit very similar time-dependent behavior. Again, this is to be expected because the only compartment that produces virus is T_1^*.

If we focus our attention on the uninfected T-cell compartments (T_1 and T_2 in Fig. 5), we see that the peak of T_1 corresponds to the valley of T_2. The fact that T_1 increases and T_2 decreases for most of the infection implies that the effects of the activation terms $p_T\left(a_T V_I/(V_I + K_V) + a_A\right) T_2$ and $\left(a_T V_I/(V_I + K_V) + a_A\right) T_2$ dominate over the differentiation terms $\gamma_T T_1$ and other loss terms in model (2.1) during the infection. That is, the T_2 compartment is losing more through activation then it is gaining through differentiation of T_1, and vice versa for T_1. The situation is reversed as the infection clears (day 50), where there is a slow decay of T_1 and corresponding rise in T_2 back toward the equilibrium values, as the decay of T_1 contributes to the rise of T_2 through differentiation.
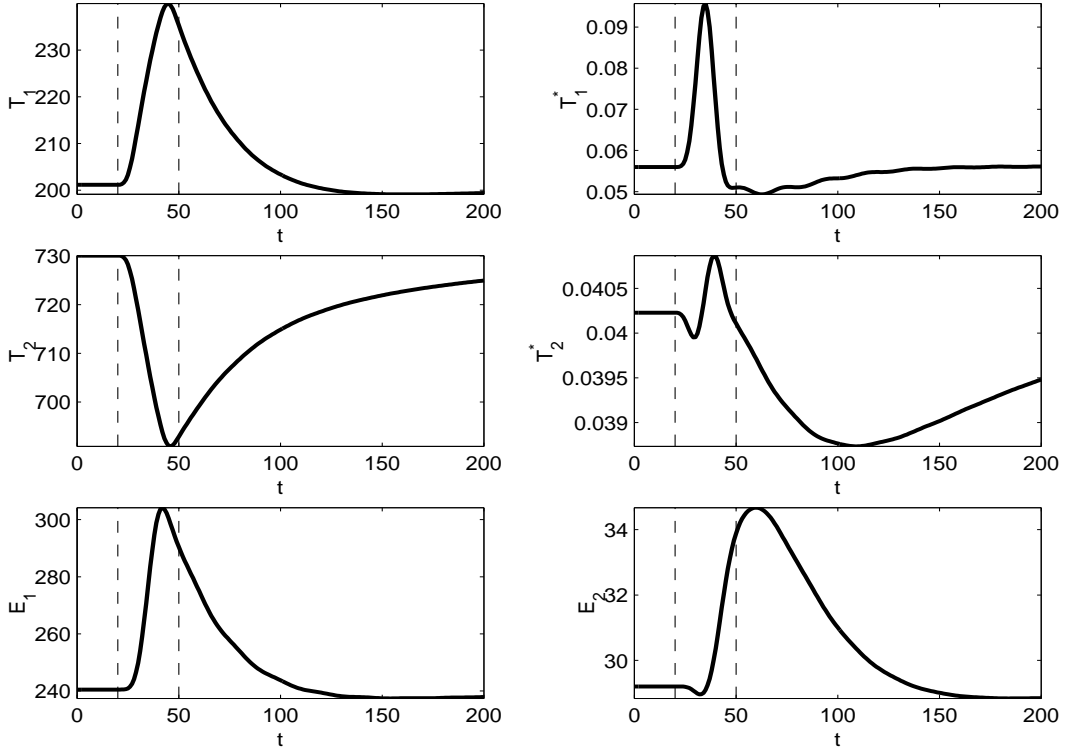
Figure 5: Model dynamics with $\epsilon_1 = 0.7$, $\epsilon_2 = 0$, and the values of all the other parameters as specified in Table 2. Activation rate by a non-HIV antigen $a_A(t)$ is as depicted in Fig. 4(a). Vertical dashed lines indicate the start and stop times of the non-HIV infection.

15

It is interesting to contrast the above behavior of T_1 and T_2 with the behavior of their infected counterparts (T_1^* and T_2^*). Unlike the T_2 compartment, which decreases throughout most of the infection, the latently-infected T_2^* compartment only decreases for a brief period, then rises and peaks near the midpoint of the infection (Fig. 5). In behavior similar to that in the T_2 compartment, the initial drop in T_2^* is due to activation of naive cells and reactivation of memory cells, although, in fact, the T_2^* compartment mostly represents infected memory cells, since the rate of infection of naive cells is very low. This activation is a source term for the T_1^* compartment and a loss term for the T_2^* compartment, but there is also an important source term for T_2^* that begins to dominate over the loss due to activation. As the T_1 compartment increases, there is an increased number of T_1 cells becoming infected and adding to the T_1^* compartment, which, in turn, leads to increased differentiation to the T_2^* compartment. The net result is, that after a short decrease, the T_2^* cells increase, even as T_1^* continues to increase, as the differentiation (source) term $\gamma_T T_1^*$ in the T_2^* compartment dominates over the activation (loss) term $\left(a_T V_I/(V_I + K_V) + a_A\right) T_2^*$ in model (2.1). The net result is that the latently-infected T_2^* compartment is supplemented, not depleted during the secondary infection. However, as a_A drops, the loss terms begin to dominate in the T_1^* and T_2^* compartments and both drop below their equilibrium values before slowly rising back to the equilibrium values. Indeed, 150 days after the end of the infection only T_1^* has returned to its equilibrium value (Fig.'s 4 and 5).

We now focus on the CD8+ compartments (E_1 and E_2), which are responsible for suppressing the viral blip. In Fig. 5 we can see that E_1 and E_2 respond to the infection with longer delays than the other compartments. This is reasonable since the activation of the CD8+ memory cells (E_2) does not depend on the non-HIV antigen a_A (these are HIV-specific memory cells) and activation will occur in response to a change in V_I. The CD8+ compartments also have a more complicated relationship to T_1 and T_1^*. In Fig. 5, we see that both compartments increase through most of the infection, although E_2 has a short period in which it decreases at the beginning of the infection. This decrease is probably due to the loss of cells due to the activation term $a_E V_I E_2/(V_I + K_V)$ in model (2.1). The subsequent increase in both CD8+ compartments indicates that the additional source term $b_{E1} T_1^* E_1/(T_1^* + K_{b1})$ in the E_1 compartment and the resulting differentiation $(\gamma_E E_1 (T_1 + T_1^*)/(T_1 + T_1^* + K_\gamma))$ to E_2 dominate over the activation (loss) term $a_E V_I E_2/(V_I + K_V)$ in the E_2 compartment. It is interesting to note that the peak value of E_2 occurs after the infection has cleared (around day 60). As the viral load drops, the loss in E_2 due to activation decreases, while the higher levels of E_1 continue to provide a source for E_2 through differentiation.

The viral blips produced with this model last about 10 days, not 2-3 weeks as estimated in the literature [15]. Some of this discrepancy may be due to the fact that we are using a higher detection limit (400 RNA copies/ml-plasma) than the experimental studies. Also, as we have seen, the dynamics of the V_I compartment closely follow those of the T_1^* compartment, which is the only source for V_I. This may explain why the the sharply rising viral load in Fig. 4 is not followed by a slower, two-phase decay as observed in [20], where there is another compartment of less-rapidly changing "chronic" cells that also produce virus. Also, in an effort to facilitate fitting to clinical data, we have combined naive, memory, activated, and helper (uninfected) T-cells in two compartments (T_1, T_2). Perhaps further subdividing the

16

CD4+ T-cell compartments or adding more compartments, such as a chronically infected T-cells pool or antigen-specific cells and non-antigen specific cells [20], would remedy this discrepancy.

Viral blips caused by sequential non-HIV infections. We next investigate the effect of sequential non-HIV infections by simulating two such back-to-back infections. The first infection occurs from days 20 through 50 and stays at its peak value (0.003 day^{-1}) from days 30 through 40, while the second infection occurs from days 80 through 110 and stays at its peak value (0.003 day^{-1}) from days 90 through 100 (Fig. 6(a)). As seen in Fig. 6(b), the first infection leads to a viral blip at exactly the same time as that in Fig. 4 and the second infection leads to a viral blip (starting at day 95) with the same duration as the first, but with a smaller amplitude.

(a)   (b)
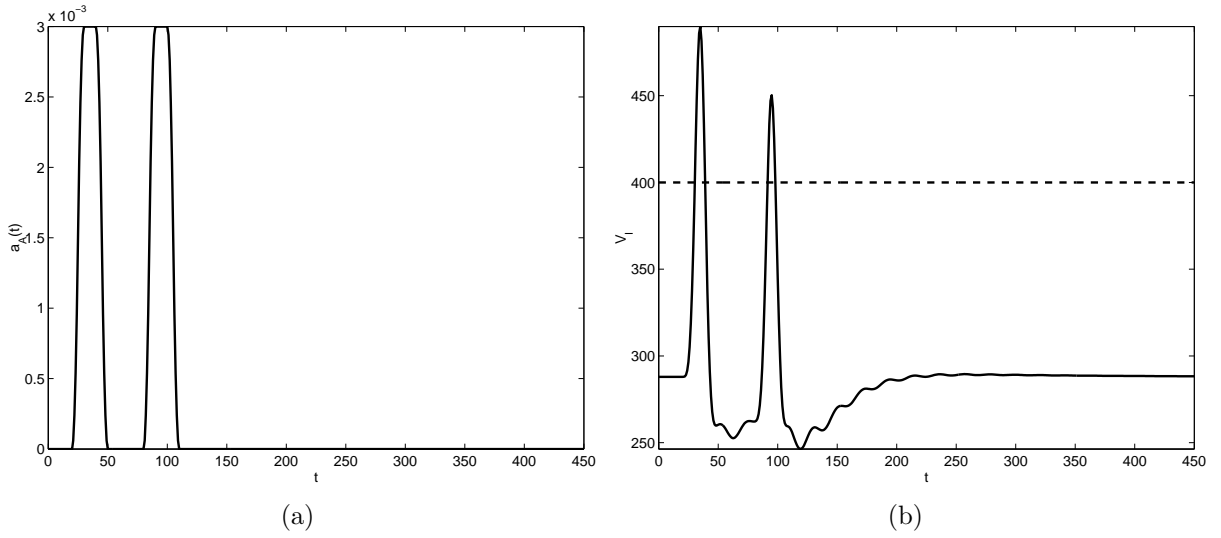
Figure 6: (a) Activation rate by non-HIV antigens $a_A(t)$. The duration of each infection is 30 days, with the first beginning on day 20 and the second on day 80; (b) Viral load $V_I(t)$ (solid line) and censor data level (horizontal dashed line).

To investigate whether the short time interval between these two infections leads to the suppressed amplitude of the second viral blip, we also considered a case where the first infection is the same as that in Fig. 6 (days 20 through 50) but the second infection occurs much later (days 200 through 230). The simulation results are illustrated in Fig. 7, which indicates that the amplitude of the second viral blip now is slightly larger than the first viral blip.

17
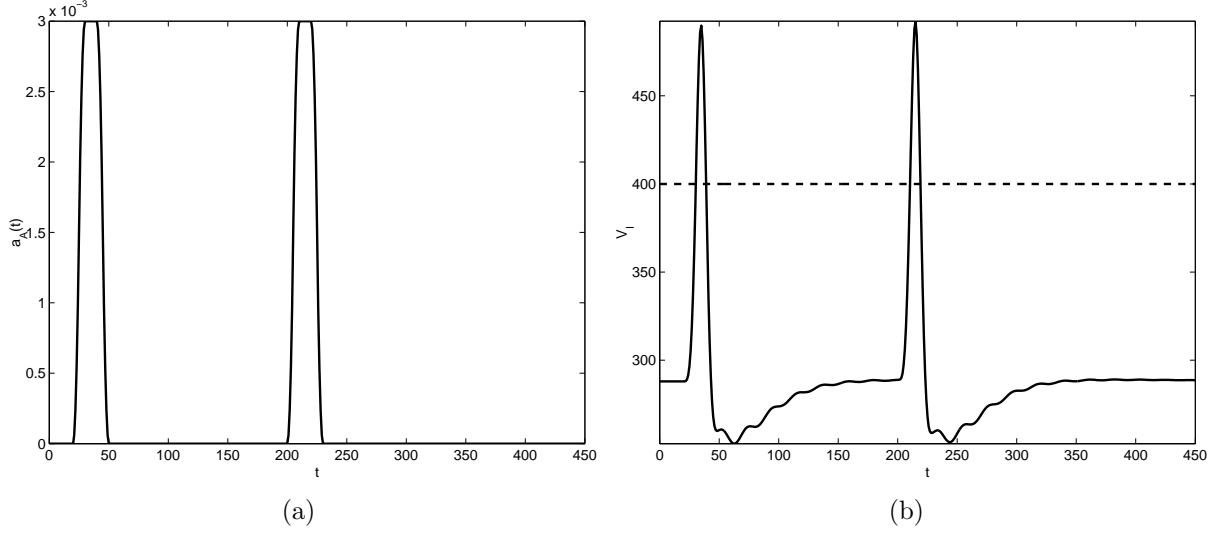
Figure 7: (a) Activation rate by non-HIV antigens $a_A(t)$. The duration of each infection is 30 days, with the first beginning on day 20 and the second on day 200; (b) Viral load $V_I(t)$ (solid line) and censor data level (horizontal dashed line).

Based on these examples, we see that the frequency of non-HIV infections affects the frequency of the viral blips and that the time interval between the two infections can affect the amplitude of the second viral blip. In particular, the shorter the time interval between two infections, the smaller the amplitude of the second viral blip. This suggests that it is possible that there may be only one viral blip detected if the time interval between the two infections is short enough.

To further investigate the dynamics of two sequential infections, we plot the other model compartments for the example shown in Fig. 7. In Fig. 8 we observe that the E_1 and E_2 compartments are still below their equilibrium values when the second infection occurs. This may explain the increased amplitude of the second viral blip. In the case of the more closely-spaced second infection of Fig. 6(a), the start of which is indicated with the dotted vertical line in Fig. 8, we can see that E_1 and E_2 are above their equilibrium levels. This may explain the decreased amplitude of the second viral blip in Fig. 6(b).

18
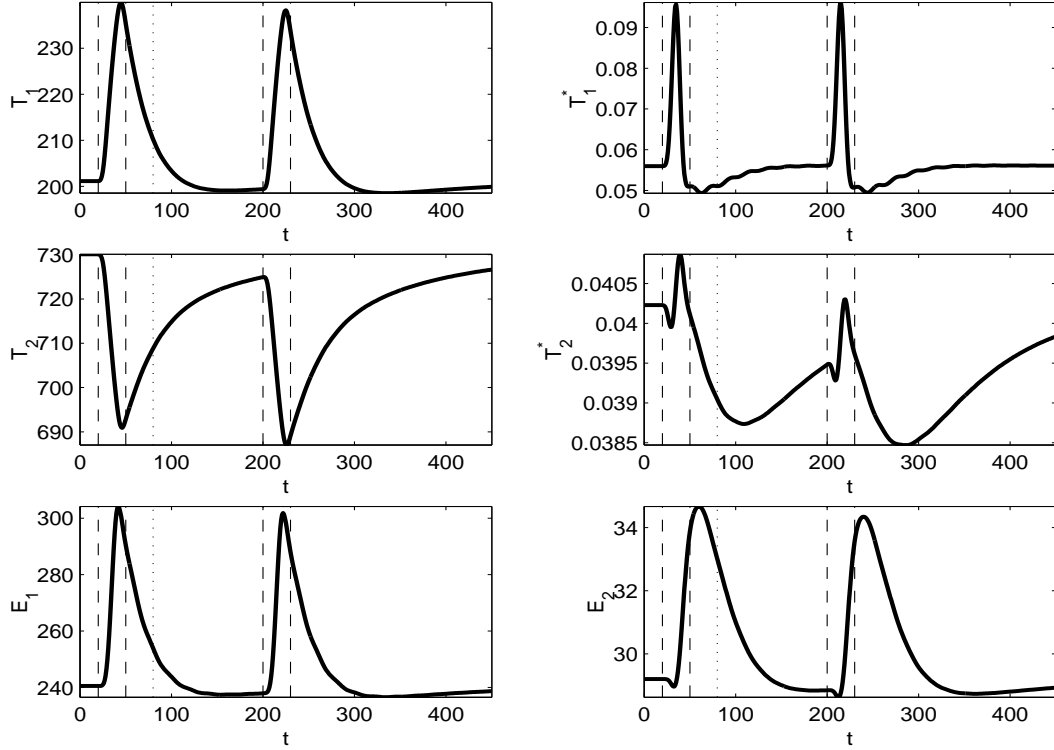
Figure 8: Model dynamics with $\epsilon_1 = 0.7$, $\epsilon_2 = 0$, and the values of all the other parameters as specified in Table 2. Activation rate by a non-HIV antigen $a_A(t)$ is as depicted in Fig. 7(a). Vertical dashed lines indicate the start and stop times of the non-HIV infections for this example. The vertical dotted line indicates the start time (80 days) of the second infection in Fig. 6.

The effect of the form of the non-HIV infection on the viral blip. As one might expect, both the shape and the peak value of $a_A(t)$ affect the amplitude and duration of the viral blip. As depicted in Fig. 9, we find that with the same shape $a_A(t)$ but with different amplitude ($a_A^{(1)}$ and $a_A^{(3)}$), the viral blip tends to have higher amplitude and a little bit wider duration as peak value of $a_A(t)$ gets higher (see $V_I^{(1)}$ and $V_I^{(3)}$ in Fig. 9(b)). Furthermore, we see that amplitude of the viral blip is lower as the non-HIV infection becomes longer and the peak value of the viral blip also occurs earlier ($V_I^{(1)}$ and $V_I^{(2)}$ in Fig. 9(b)).
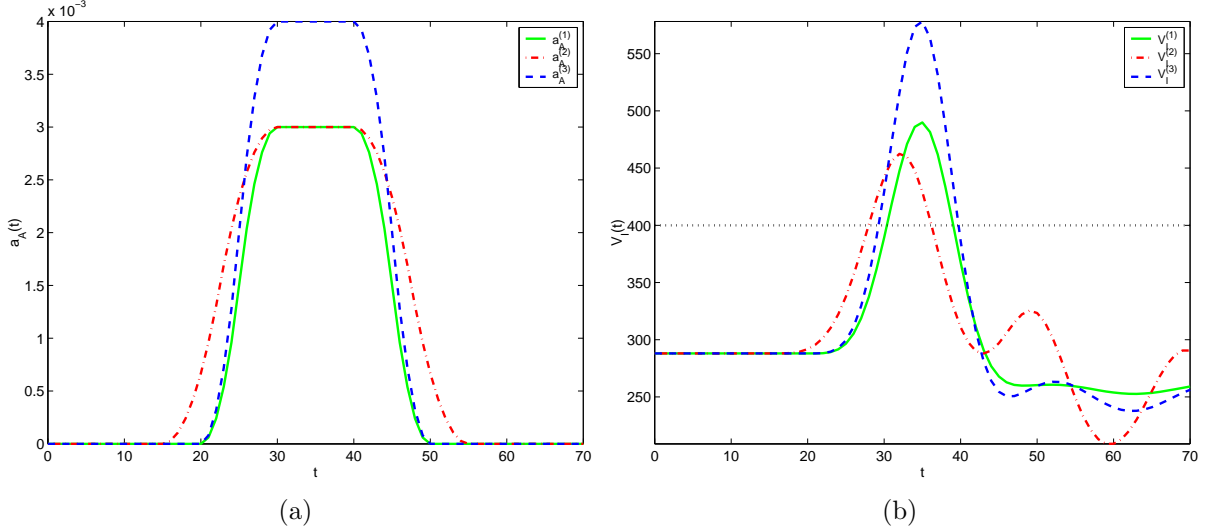
Figure 9: (a) Activation rate by non-HIV antigen $a_A(t)$; (b)Viral load $V_I(t)$.

CD4+ help during a viral blip. In model (2.1), the differentiation of E_1 cells to E_2 memory cells occurs through the term $\gamma_E \frac{T_1+T_1^*}{T_1+T_1^*+K_\gamma}E_1$, which is dependent on the help of activated CD4+ cells (T_1 and T_1^*). In the case that T_1 and T_1^* are zero, there is no differentiation of the effector E_1 cells into memory E_2 cells. On the other hand, in the limit of large $T_1 + T_1^*$, the maximum rate of differentiation $\gamma_E E_1$ is attained. In this section we explore what effect the CD4+ help has on the CD8+ immune response and we do so in the context of a viral blip induced by a non-HIV infection $a_A(t)$. We use the parameter K_γ to modify the effects of the CD4+ help. The case of $K_\gamma = 0$ simulates a condition where CD4+ help is not required for differentiation to CD8+ memory, while the case of K_γ very large simulates a condition where differentiation to CD8+ memory is impaired due to lack of CD4+ help. Thus, we use the states of locally a.s. "healthy" infected equilibria, corresponding to $\epsilon_1 = 0.7$, $\epsilon_2 = 0$, $a_A = 0$, various values of K_γ, and the values of all other parameters as specified in Table 2, as the initial conditions for simulations in which we allow the parameter a_A to be a function of time t. As we drive the system (2.1) with the non-HIV infection $a_A(t)$ we monitor the states of the system.

In Fig. 10(a) we plot the "healthy" infected (EQ$_2$) equilibrium values of the V_I, E_1, and E_2 compartments as a function of K_γ. As expected, as K_γ increases, corresponding to increasing impairment of CD8+ differentiation to memory, the equilibrium value of the E_2

20

compartment decreases. However, the overall effect on E_1 is negligible with only a slight rise as K_γ increases. Despite the fact that E_1 stays relatively constant, we see that, as K_γ increases, the equilibrium value of the viral load increases until reaching a plateau value at $K_\gamma \approx 1 \times 10^5$ cells/μl-blood. Thus, even though the number of effector cells, which directly participate in the removal of infected T-cells, does not decrease as the CD8+ differentiation to memory is impaired, the overall viral load increases.
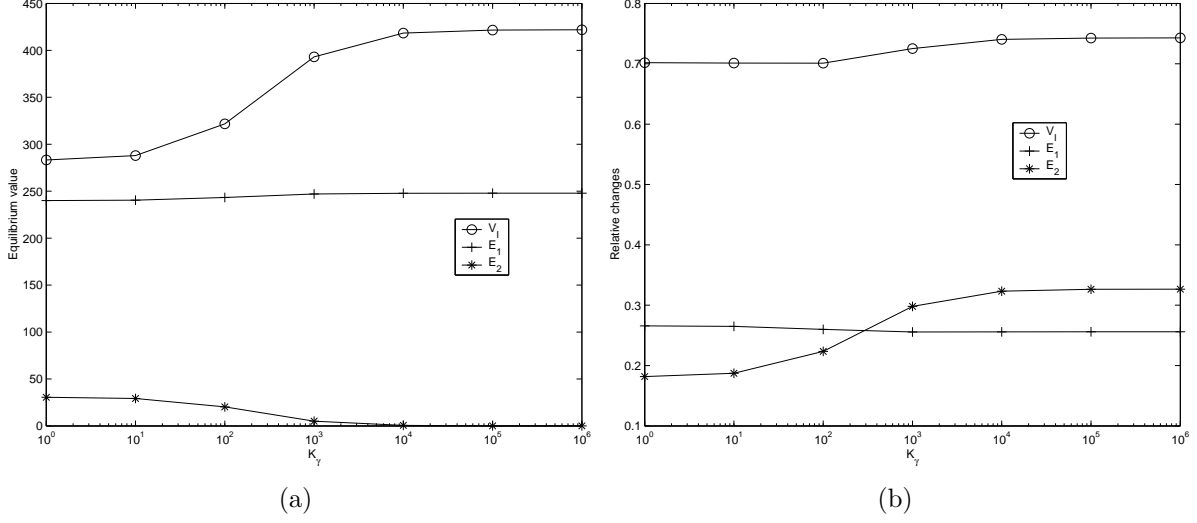
Figure 10: (a) The effect of K_γ on the "healthy" infected (EQ$_2$) values of the V_I (circles), E_1 (crosses), and E_2 (asterisks) compartments. Equilibrium values correspond to the case $\epsilon_1 = 0.7$, $\epsilon_2 = 0$, $a_A = 0$, and the values of all other parameters as specified in Table 5; (b) Relative change of V_I (circles), E_1 (crosses), and E_2 (asterisks) compartments as a function of K_γ.

In Fig. 10(b) we plot the relative changes of the various compartments as a function of K_γ. Relative change is defined as (peak value minus initial value)/(initial value). In this figure we can see that, as K_γ increases the relative change of the peak viral load increases. The relative change of the E_2 memory compartment also increases (Fig. 10(b)), but the infected equilibrium value was quite small to begin with (Fig. 10(a)). The relative change in the effector E_1 compartment decreases as K_γ increases. The results of this analysis demonstrates that in our model, impaired differentiation to CD8+ memory results in a degraded "healthy" equilibrium and leads to a larger viral blip following a secondary infection.

5   Fitting the Model to Clinical Data

When proposing a new model, a necessary but often difficult requirement is validation of the model with clinical data. One way to do this is to use the first part of the available longitudinal data for a given patient to estimate patient specific parameters. One then uses these parameters in model simulations to accurately predict the subsequent progression of the disease as represented in the remainder of the longitudinal data for that patient. In this section we test the capability of model (2.1) to fit the clinical data of patients and to predict the viral progression. Of course, the dynamics of HIV progression may totally change as time evolves (e.g., such as transfer from "unhealthy" steady state region of attraction to one for a "healthy" state). In this section, we test the predictive capabilities of the model with parameters that have been estimated from only *half* of the longitudinal data by comparing simulations to the full set of clinical observations.

To obtain patient specific-parameter estimates for the model, we used individual patient data (either partial or full longitudinal data) and carried out an inverse problem. The data for our investigations come from Massachusetts General Hospital (MGH), where all the patients enrolled in the study are symptomatic with acute or early HIV-infection (for more detailed information of these data, the interested reader is referred to [3, 21, 27]). In summary, nearly all subjects in the study undergo combination therapy and many have at least one treatment interruption. The available clinical data include total CD4+ T-cell count and total RNA copies, where for model (2.1) the total CD4+ T-cell counts are represented by $\bar{z}_1(t; q) = T_1(t; q) + T_1^*(t; q) + T_2(t; q) + T_2^*(t; q)$ and total RNA copies are represented by $\bar{z}_2(t; q) = V_I(t; q) + V_{NI}(t; q)$. If the measurements of RNA copies are below the limit of quantification for the assay used (400 copies/ml-plasma for a standard assay and 50 copies/ml-plasma for an ultra-sensitive assay), then the observed viral load value is censored to be at its detection limit; that is, in these cases the observed values do not represent the true data values anymore. Furthermore, observations of viral load and CD4+ may not be at the same time points and the observation times and intervals vary substantially among patients. So, in general, for patient number j we have CD4+ T-cell data pairs $(t_1^{ij}, \bar{y}_1^{ij})$, $i = 1, \cdots, N_1^j$, and potentially different time point viral RNA data pairs $(t_2^{ij}, \bar{y}_2^{ij})$, $i = 1, \cdots, N_2^j$. Hence, the clinical data for carrying out the inverse problem involves *partial observations*, *measurements from combined compartments*, and *highly censored viral load measurements*.

As one might expect, if a patient does not have a sufficient number of observations or does not undergo a therapy interruption during the observation period used to fit the model, then it is difficult to obtain dynamically dependent parameters to use to predict a later disease progression involving both therapy and treatment interruption. Hence, for our purposes we focused on 14 patients wherein each patient has ten or more each of CD4+ and viral load measurements, and have at least one on/off treatment schedule in the first half of their longitudinal data. Moreover, for each of these patients, N_1^j is not substantially different than N_2^j. The technique we use in this paper is adapted from the one in [3], where the authors developed a statistically-based censored data method (an expectation maximization algorithm) combined with an ordinary nonlinear least-squares technique. We note that

the variance σ_1^2 in CD4 measurements and the variance σ_2^2 in viral load measurements are likely to be different due to assay differences. Hence, in this paper we use the expectation maximization algorithm based on Maximum Likelihood Estimation for $(q, \sigma_1^2, \sigma_2^2)$ which, under normality assumptions on the errors, results in a weighted least-squares technique (see [30]) with solution given by

$$
\hat{q}^j = \arg\min_{q \in Q} \left[\frac{1}{\hat{\sigma}_1^2} \sum_{i=1}^{N_1^j} |y_1^{ij} - z_1(t_1^{ij}; q)|^2 + \frac{1}{\hat{\sigma}_2^2} \sum_{i=1}^{N_2^j} |y_2^{ij} - z_2(t_2^{ij}; q)|^2 \right]
$$
(5.1)
$$
\hat{\sigma}_k^2 = \frac{1}{N_k^j} \sum_{i=1}^{N_k^j} |y_k^{ij} - z_k(t_k^{ij}; \hat{q}^j)|^2, \quad k = 1, 2,
$$

for the $\log_{10}$-transformed system of model (2.1) for patient j, where $y_1^{ij} = \log_{10} \bar{y}_1^{ij}$ and $z_1(t_1^{ij}; q) = \log_{10}(\bar{z}_1(t_1^{ij}; q))$, $i = 1, \cdots, N_1^j$, and $y_2^{ij} = \log_{10} \bar{y}_2^{ij}$, $z_2(t_2^{ij}; q) = \log_{10}(\bar{z}_2(t_2^{ij}; q))$, $i = 1, \cdots, N_2^j$. As noted in [3], by using a log-transformed system one can resolve a problem of states becoming unrealistically negative due to round-off error: nonnegative solutions of this model should stay so throughout numerical simulation. This approach also enables efficient handling of unrealistic cases where states get infinitesimally small during integration due to parameters selected by optimization algorithms. From a statistical point of view, log transformation is a standard technique to render the observations more nearly normally distributed, which also supports use of the weighted least squares criterion as an equivalent to maximum likelihood estimation.

The expectation maximization (EM) algorithm is outlined below. To simplify the notation, we drop the patient index j in this algorithm description. The following notation will be used in the algorithm: the relevant censoring point at time t^i is represented by L^i and χ^i is the indicator function for the set $\{y_2^i > L^i\}$, ϕ denotes the standard normal probability density function and Φ is the corresponding cumulative distribution function. For each patient, we carried out the following parameter estimation algorithm:

- (Step 1) Create adjusted data $\tilde{y}^i$ by replacing censored y_2^i values by $L^i/2$, and use ordinary least squares to estimate $\hat{q}^{(0)}$ using both CD4+ data y_1^i and adjusted viral RNA data $\tilde{y}^i$ (which now includes replaced censored values).

$$
\hat{q}^{(0)} = \arg\min_{q \in Q} \left[\sum_{i=1}^{N_1} |y_1^i - z_1(t_1^i; q)|^2 + \sum_{i=1}^{N_2} |\tilde{y}^i - z_2(t_2^i; q)|^2 \right].
$$

Obtain an initial estimate for σ_1^2 and σ_2^2 from

$$
\left(\hat{\sigma}_1^{(0)} \right)^2 = \frac{1}{N_1} \sum_{i=1}^{N_1} |y_1^i - z_1(t_1^i; \hat{q}^{(0)})|^2, \quad \left(\hat{\sigma}_2^{(0)} \right)^2 = \frac{1}{N_2} \sum_{i=1}^{N_2} |\tilde{y}^i - z_2(t_2^i; \hat{q}^{(0)})|^2,
$$

respectively. Set $k = 0$.

23

- (Step 2) Define $\hat{z}_2^{i(k)} = z_2(t_2^i; \hat{q}^{(k)})$ and $\hat{\zeta}^{i(k)} = \frac{L^i - \hat{z}_2^{i(k)}}{\hat{\sigma}_2^{(k)}}$. Update the data and residuals by

$$\begin{aligned}
\tilde{y}^{i(k)} &= \chi^i y_2^i + (1 - \chi^i)\left[\hat{z}_2^{i(k)} - \hat{\sigma}_2^{(k)} \frac{\phi(\hat{\zeta}^{i(k)})}{\Phi(\hat{\zeta}^{i(k)})}\right], \\
\tilde{r}^{i(k)} &= \chi^i(y_2^i - \hat{z}_2^{i(k)})^2 + (1 - \chi^i)(\hat{\sigma}_2^{(k)})^2\left[1 - \hat{\zeta}^{i(k)}\frac{\phi(\hat{\zeta}^{i(k)})}{\Phi(\hat{\zeta}^{i(k)})}\right].
\end{aligned}$$

- (Step 3) Update the estimate of q to $\hat{q}^{(k+1)}$ by performing the weighted least squares minimization in the parameters q

$$\hat{q}^{(k+1)} = \arg\min_{q \in Q}\left[\frac{1}{(\hat{\sigma}_1^{(k)})^2}\sum_{i=1}^{N_1}|y_1^i - z_1(t_1^i; q)|^2 + \frac{1}{(\hat{\sigma}_2^{(k)})^2}\sum_{i=1}^{N_2}|\tilde{y}^{i(k)} - z_2(t_2^i; q)|^2\right],$$

and computing

$$\left(\hat{\sigma}_1^{(k+1)}\right)^2 = \frac{1}{N_1}\sum_{i=1}^{N_1}|y_1^i - z_1(t_1^i; \hat{q}^{(k+1)})|^2, \quad \left(\hat{\sigma}_2^{(k+1)}\right)^2 = \frac{1}{N_2}\sum_{i=1}^{N_2}\tilde{r}^{i(k)}.$$

If relative changes in $\hat{q}$, $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are small, terminate. Otherwise set $k = k + 1$ and then go to Step 2.

For more details about the EM algorithm and how to carry it out, interested readers are referred to [3, 13, 23].

Note that model (2.1) has 31 model parameters and 8 initial conditions. The total number of measurements (sum of number of measurements of CD4+ and number of measurements of RNA copies) in the first half of the longitudinal data varies from 42 to 154 in these 14 patients, which means it is difficult to estimate all 39 of these parameter values for some of these patients by using half of the longitudinal data set. Hence, we first try to estimate all the 31 model parameters and 8 initial conditions for each of the 14 patients by applying the EM algorithm to the full longitudinal data set. We then try to fix some model parameters and initial conditions at the *population averages* across these patients, and attempt to estimate all the remaining model parameters and initial conditions for each patient by applying the EM algorithm separately to each of the first half and the full longitudinal data set. Note that there exists biological variation in all parameters across the patients, and there also exist high correlations among some of these parameters such as the RNA copies produced per infected cells N_T and the virus natural death rate c. We also observe that sensitivity with respect to some of these parameters may be highly time-dependent. For example, the dynamics of the model is much more sensitive to the treatment efficacies ϵ_1 and ϵ_2 in the treatment periods than it is in the off-treatment periods. All of these considerations make it difficult to choose a priori which parameters can be fixed. In this paper, we empirically chose some parameters (such as the saturation parameters) to which the model appears to be relatively insensitive to take as fixed. Table 5 specifies all the fixed parameters and their corresponding values. After we obtain the two sets of model parameters (corresponding to

γ_T	3.792e-04	d_2	3.096e-03	f	5.068e-01
k_2	2.005e-09	δ	2.095e-01	m	1.127e-03
c	5.818e+00	λ_E	9.930e-04	b_{E1}	3.885e-02
K_{b1}	2.488e-02	d_E	6.278e-02	K_d	1.200e-01
δ_{E1}	5.967e-02	K_{b2}	8.697e+01	γ_E	5.154e-04
K_γ	1.357e+00	K_V	1.479e+01	δ_{E2}	1.450e-03
K_s	2.789e+04	T_2^{*0}	7.521e-03	V_{NI}^0	3.571e+03
E_1^0	6.821e-02	E_2^0	6.909e-01		

Table 5: Average model parameter values (19) and initial conditions (4) used in model fitting with half and full longitudinal data set.

use of the first half of the data and the full set of data, respectively, in inverse problems), we simulate the trajectory over the full time span of the patient's observations by using the parameter values obtained with these two data sets and compare their ability to describe the experimental results.

Figures 11, 12, and Fig.'s 13-24 in the Appendix illustrate model fits for all 14 patients. These figures reveal that both the fits to CD4+ T-cells and the fits to viral load are reasonable when using either half data or full data to estimate the parameter values. However, there is a small amount of under-prediction or over-prediction for some patients when using parameters estimated from the half data set. For example, model fits obtained by using half and full time series data for patient 2 are shown in Fig. 11, where it can be seen that by using parameters
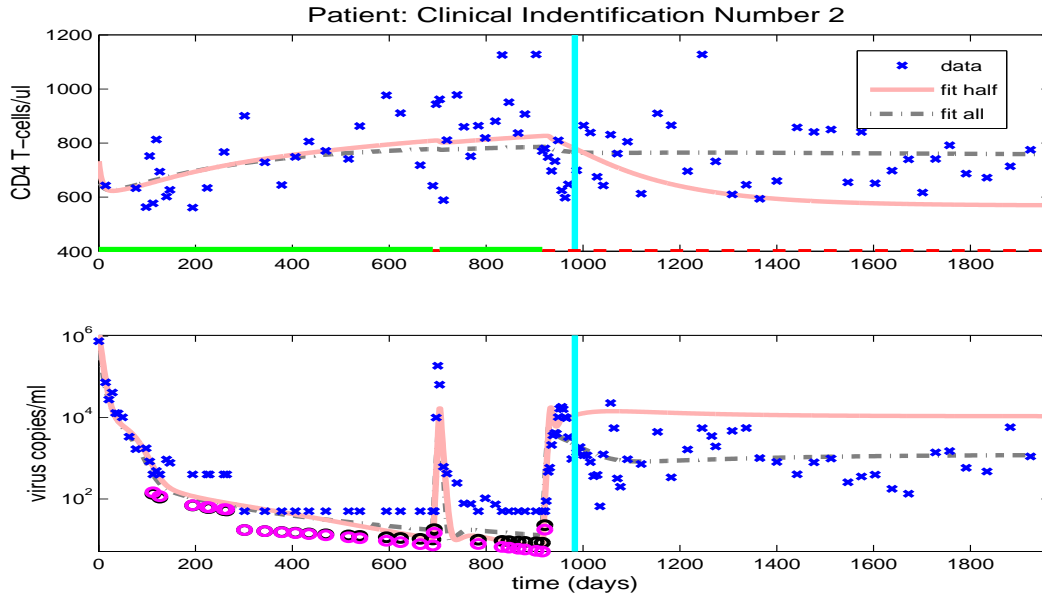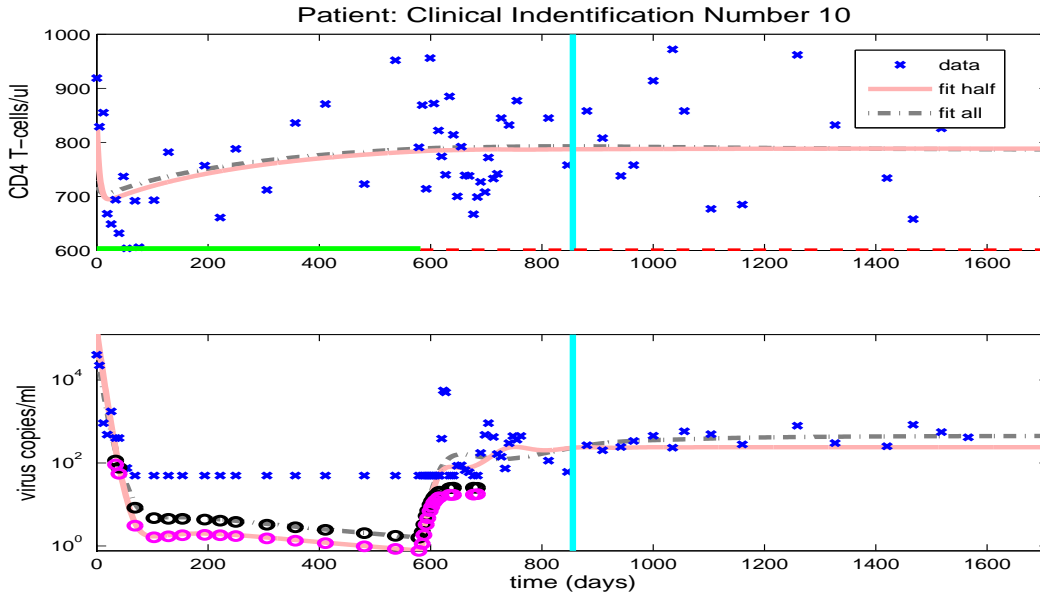

Figure 11: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.

estimated from the first half data set, the predicted values for viral load data are slightly

higher and the predicated values for CD4+ T-cell are slightly lower than the observed data values and the fitted values from using parameters estimated from the full data, even though there are two off-treatment periods in the first half of the data set. This discrepancy may be due to the fact that both off-treatment periods are short and, therefore, provide very little information about the off-treatment behavior. Note that the values of most observations of viral loads in the off-treatment periods during the first half data set are higher than those observed in the second half data set and this makes it difficult to correctly predict future off-treatment trends. Hence, it is not just the number of off-treatment periods in the first half data set but also the length of each off-treatment period and the number of observations in these periods that determines the accuracy of the predictions of off-treatment viral load levels.

Figure 12 compares the model fit and prediction obtained by using the half data set to the fit obtained from parameter estimation with full data set for patient 10. Even though there is only one off-treatment period in the first half data set, the period is sufficiently long and the data set sufficiently rich that the full data fit and half data predictions of both CD4+ T-cell and viral load agree quite well in the second half data set. Hence, the results demonstrated in Fig.'s 11 and 12 suggest that in order to have accurate predictive capability we need to obtain a sufficient number of representative data points in the time period providing the information about the dynamics (the time period used to estimate parameters). Finally, we note that the model developed and validated here retains the predictive capabilities present in the earlier model of [3] in that the ability to predict subsequent data after fitting the model with half of the data is at least as good as that exhibited in [3] for the various patients in our data sets.

Figure 12: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.

26

In Section 3 we demonstrated that model (2.1) with a specific example set of parameters exhibited multiple equilibria. Similarly, after fitting the full set of patient clinical data we used the parameter sets thus obtained to determine the off-treatment locally a.s. equilibria for each patient, and we find that, for all but two of the patients, there is only one locally a.s. equilibrium available for each patient. Obviously, these equilibria are all infected equilibria, with equilibrium viral loads ranging from 450.2 to 51623 copies/ml-plasma. In contrast, patients 25 and 26 each have one locally a.s. uninfected equilibrium and one locally a.s. infected equilibrium, although the uninfected equilibrium for both these cases exhibit physiologically unreasonable levels of HIV-specific CD8+ cells as compared to the infected equilibrium values. This behavior is not unusual since we do not, at this point, have any clinical data for HIV-specific CD8+ cell counts to help constrain these compartments. In future efforts with data from clinical trials currently being designed, we plan to explore the use of an additional cost term in the optimization algorithm that constrains the CD8+ values to fit observations.

6 Conclusion and Remarks

One of the challenges in our clinical data fitting is that we do not have sufficiently luxurious data sets to fit all the model parameters and initial conditions. Due to the high correlations between the model parameters, biological variations in all parameters across population and the sensitivity with respect to parameters varying over time (especially during the transition time between the off-treatment period and on-treatment period), we intuitively chose some parameters in this paper to be fixed. Even though the simulation results indicate reasonable fits to all the patients, we need to develop a scientific methodology to deal with this situation in order to obtain more reliable parameter values to be fixed a priori and thereby sharpen our estimation results. As outlined above we reduced the number of parameters to be estimated by fixing some of the model parameters at population averages. In clinical cases where large, long term individual patient longitudinal data sets are not available as in our study here, one could use population averages accrued in prior patient trials to reduce the number of free parameters. Another possible method to reduce the dimension of parameter space and the associated high correlation between these parameters is principal component analysis, which is currently under investigation.

The clinical data fitting results in Section 5 indicate that if one does not have a sufficient number of observations during the periods where the dynamics are changing, then it is difficult to obtain parameter estimates useful in predicting corresponding future trends in disease progression. Hence, the goodness of fit results are affected by not only the *number* of observations but also the sampling *times* for data collection. This general principle was also observed and explained conceptually in [6] where the general "information content" in data and its relation to sensitivity is explored. These considerations suggest that parameter sensitivity with respect to data is important in experimental design, assisting in reducing effort and resources required to collect necessary data. One potential approach for this is the generalized sensitivity function methodology as proposed in [36] and explored in [5, 7]. This

approach combines the sensitivities of model output with respect to model parameters with the sensitivities of parameters estimates with respect to changes in model outputs. Obtaining results using this methodology is challenging even for simple examples and our initial efforts have not yet proved fruitful for our complex HIV models which feature high correlations between some parameters and dynamics that dramatically change during the transition time between off-treatment and on-treatment. Hence, one of our future efforts involves development of a methodology to determine improved sampling times for data collection. We are optimistic that ideas contained in principal component analysis and generalized sensitivity functions can be combined to provide new guidance in this regard.

Acknowledgements

References

[1] B.M. Adams, H.T. Banks, M. Davidian, H. Kwon, H.T. Tran, S.N. Wynne and E.S. Rosenberg, HIV dynamics: modeling, data analysis, and optimal treatment protocols, *J. Computational and Applied Mathematics*, **184** (2005), 10–49.

[2] B.M. Adams, H.T. Banks, H.T. Tran and H. Kwon, Dynamic multidrug therapies for HIV: optimal and STI control approaches, *Mathematical Biosciences and Engineering*, **1** (2004), 223–241.

[3] B.M. Adams, H.T. Banks, M. Davidian and E.S. Rosenberg, Model fitting and prediction with HIV treatment interruption data, CRSC-TR05-40, NCSU, October, 2005; *Bulletin of Mathematical Biology*, **69** (2007), 563–584.

[4] M. Altfeld and E.S. Rosenberg, The role of CD4 T helper cells in the cytotoxic T lymphocyte response to HIV-1, *Current Opinion in Immunology*, **12** (2000), 375–380.

[5] P. Bai, H.T. Banks, S. Dediu, A.Y. Govan, M. Last, A. Loyd, H.K. Nguyen, M.S. Olufsen, G. Rempala and B.D. Slenning, Stochastic and deterministic models for agricultural production networks, CRSC-TR07-06, NCSU, February, 2007; *Mathematical Biosciences and Engineering*, to appear.

[6] H.T. Banks, S.L. Ernstberger and S.L. Grove, Standard errors and confidence intervals in inverse problems: Sensitivity and associated pitfalls, CRSC-TR06-10, NCSU, March, 2006; *J. Inverse and Ill-posed Problems*, to appear.

[7] J. J. Batzel, F. Kappel, D. Schneditz and H.T. Tran, *Cardiovascular and Respiratory Systems: Modeling, Analysis and Control*, SIAM Frontiers in Applied Math, SIAM, Philadelphia, 2006.

[8] M.J. Bevan, Helping the CD8+ T-cell response, *Nature Reviews*, **4** (2004), 595–602.

[9] J.N. Blankson, D. Persaud, and R.R. Siliciano, The challenge of viral reservoirs in HIV-1 infection, *Annual Review of Medicine*, **53** (2002), 557–593.

[10] S. Bonhoeffer, M. Rembiszewski, G.M. Ortiz and D.F. Nixon, Risks and benefits of structured antiretroviral drug therapy interruptions in HIV-1 infection, *AIDS*, **14** (2000), 2313–2322.

[11] J.M. Brenchley, B.J. Hill, D.R. Ambrozak, D.A. Price, F.J. Guenaga, J.P. Casazza, J. Kuruppu, J. Yazdani, S.A. Migueles, M. Connors, M. Roederer, D.C. Douek and R.A. Koup, T-cell subsets that harbor human immunodificiency virus (HIV) in vivo: implications for HIV pathogenesis, *J. of Virology*, **78** (2004), 1160–1168.

[12] D.S. Callaway and A.S. Perelson, HIV-1 infection and low steady state viral loads, *Bulletin of Mathematical Biology*, **64** (2001) 29–64.

[13] A.P. Dempster, N.M. Laird and D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Stat. Soc., Ser. B*, **39** (1977),1–38.

[14] T-W. Chun, L. Carruth, D. Finzi, X.F. Shen, J.A. DiGiuseppe, H. Taylor, M. Hermankov, K. Chadwick, J. Margolick, T.C. Quinn, Y.H. Kuo, R. Brookmeyer, M.A. Zeiger, P. BarditchCrovo and R.F. Siliciano, Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection, *Nature*, **387** (1997), 183–188.

[15] M. Di Mascio, M. Markowitz, M. Louie, C. Hogan, A. Hurly, D.D. Ho and A.S. Perelson, Viral blip dynamics during highly active antiretroviral therapy, *J. of Virology*, **77** (2003), 12165–12172.

[16] N.M. Dixit and A.S. Perelson, Complex patternsof viral load decay under antiretroviral therapy: influence of pharmacokineticsand intracellular delay, *J. of Theoretical Biology*, **226** (2004) 95–109.

[17] D. Finzi, M. Hermankova, T. Pierson, L.M. Carruth, C. Buck, R.E. Chaisson, T.C. Quinn, K. Chadwick, J. Margolick, R. Brookmeyer, J. Gallant, M. Markowitz, D.D. Ho, D.D. Richman and R.F. Siliciano, Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy, *Science*, **278** (1997), 1295–1300.

[18] D.D. Ho, A.U. Neumann, A.S. Perelson, W. Chen, J.M. Leonard, and M. Markowitz, Rapid turnoverof plasma virions and CD4 lymphocytes in HIV-1 infection, *Nature*, **373** (1995), 123–126.

[19] E.M. Janssen, E.E. Lemmens, T. Wolfe, U. Christen, M.G. Von Herrath and S. P. Schoenberger, CD4+ T-cells are required for secondary expansion and memory in CD8+ T lymphocytes, *Nature*, **421** (2003), 852–856.

[20] L.E. Jones and A.S. Perelson, Opportunistic infection as a cause of transient viremia in chronically infected HIV patients under treatment with HAART, *Bulletin of Mathematical Biology*, **67** (2005), 1227–1251.

[21] S. Kassutto, K. Maghsoudi, M. Johnston, G. Robbins, N. Burgett, P. Sax, D. Cohen, E. Pae, B. Davis, K. Zachary, N. Basgoz, E.M.C. D'agata, V. DeGruttola, B. Walker and E. Rosenberg, Longitudinal analysis of clinical markers following antiretroviral therapy initiated during acute or early HIV-1 infection, *Clinical Infectious Disease*, **42** (2006), 1024–1031.

[22] J.M. McCune, The dynamics of CD4+ T-cell depletion in HIV disease, *Nature*, **410** (2001), 974–979.

[23] G.J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, John Wiley, New York, 1997.

[24] M.A. Nowak and C.R.M. Bangham, Population dynamics of immune responses to persistent viruses, *Science*, **272** (1996), 74–79.

[25] A.S. Perelson, P. Essunger, Y.Z. Cao, M. Vesanen, A. Hurley, K. Saksela, M. Markowitz and D.D. Ho, Decay characteristics of HIV-1-infected compartments during combination therapy, *Nature*, **387** (1997), 187–191.

[26] A.S. Perelson and P.W. Nelson, Mathematical analysis of HIV-1 dynamics in vivo, *SIAM Review*, **41** (1999), 3–44.

[27] E.S. Rosenberg, M. Altfield, S.H. Poon, M.N. Phillips, B. Wilkes, R.L. Eldridge, G.K. Robbins, R.D. D'Aquila, P.J.R. Goulder and B.D. Walker, Immune control of HIV-1 after early treatment of acute infection, *Nature*, **407** (2000),523–526.

[28] S.L. Rowland-Jones, T. Dong, K.R. Fowke, J. Kimani, P. Krausa, H. Newell, T. Blanchard, K. Ariyoshi, J. Oyugi, E. Ngugi, J. Bwayo, K.S. MacDonald, A.J. McMichael and F.A. Plummer, Cytotoxic T-cell responses to multiple conserved HIV epitopes in HIV-resistant prostitute in Nairobi, *J. of Clinical Investigation*, **102** (1998), 1758–1765.

[29] L.A. Pinto, J. Sullivan, J.A. Berzofsky, M. Clerici, H.A. Kessler, A.L. Landay and G.M. Shearer, ENV-specific cytotoxic T lymphocyte responses in HIV seronegative health care workers occupationally exposed to HIV-contaminated body fluids, *J. of Clinical Investigation*, **96** (1995), 867–876.

[30] G. A. F. Seber and C. J. Wild, *Nonlinear Regression,* John Wiley & Sons, Inc., New York, 1989.

[31] R.A. Seder and R. Ahmed, Similarities and differences in CD4+ and CD8+ effector and memory T-cell generation, *Nature Immunology*, **4** (2003), 835–842.

[32] D.J. Shedlock and H. Shen, Requirement for CD4 T-cell help in generating functional CD8 T-cell memory, *Science*, **300** (2003), 337–339.

[33] J. D. Siliciano, J. Kajdas, D. Finzi, T. C. Quinn, K. Chadwick, J. B. Margolick, C. Kovacs, S. J. Gange, and R. F. Siliciano, Long-term follow-up studies confirmed the stability of the latent reservoir for HIV-1 in resting CD4+ T-cells, *Nature Medicine*, **9** (2003), 727–728.

[34] J. Sprent and C. D. Surh, T-cell memory, *Annual Rev. Immunol.*, **20** (2002), 551–579.

[35] H. Streeck, H. Jessen, G. Alter, N. Teigen, M.T. Waring, A. Jessen, I. Stahmer, J. van Lunzen, M. Ichterfeld, X.J. Gao, T.M. Allen, M. Carrington, B.D. Walker, J.K. Rockstroh and M. Altfeld, Immunological and virological impact of highly active antiretroviral therapy initiated during acute HIV-1 infection, *J. of Infectious Diseases*, **194** (2006), 734-739.

[36] K. Thomaseth and C. Cobelli, Generalized sensitivity functions in physiological system identification, *Annals of Biomedical Engineering*, **27** (1999), 607–616.

[37] C. Utzny and N.J. Burroughs, Long-term stability of diverse immunological memory, *J. Theor. Biol.*, **211** (2001), 393–402.

[38] R. M. Welsh, L. K. Selin and E. Szomolanyi-Tsuda, Immunological memory to viral infections, *Annual Rev. Immunol.*, **22** (2004), 711–743.

[39] D. Wodarz and M.A. Nowak, Specific therapy regimes could lead to long-term immunological control of HIV, *Proceedings National Academy of Sciences*, **96** (1999), 14464–14469.

Appendix: Section 5 Results for Other Patients
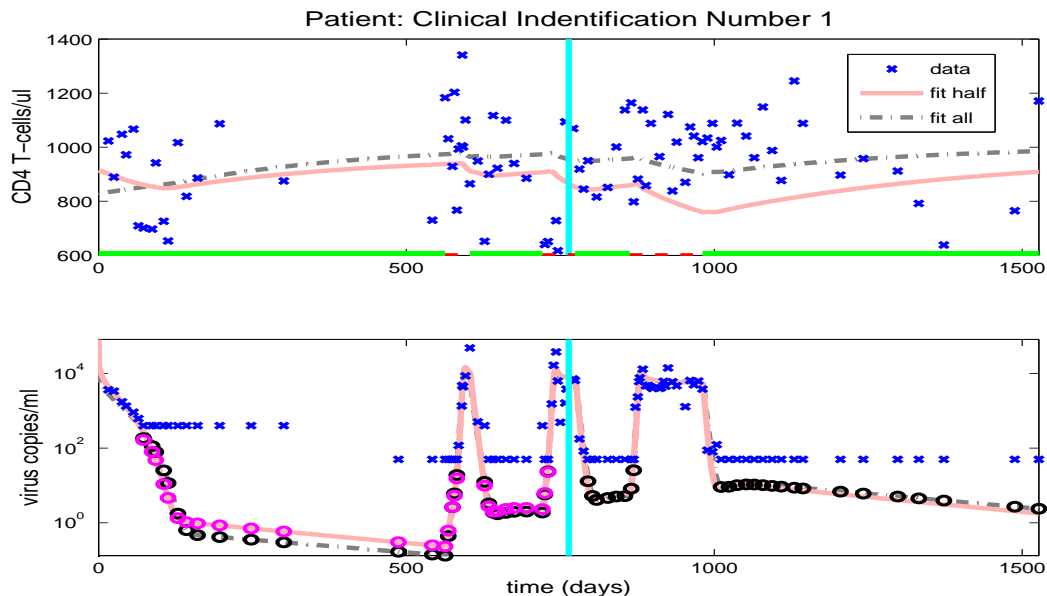
Figure 13: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.
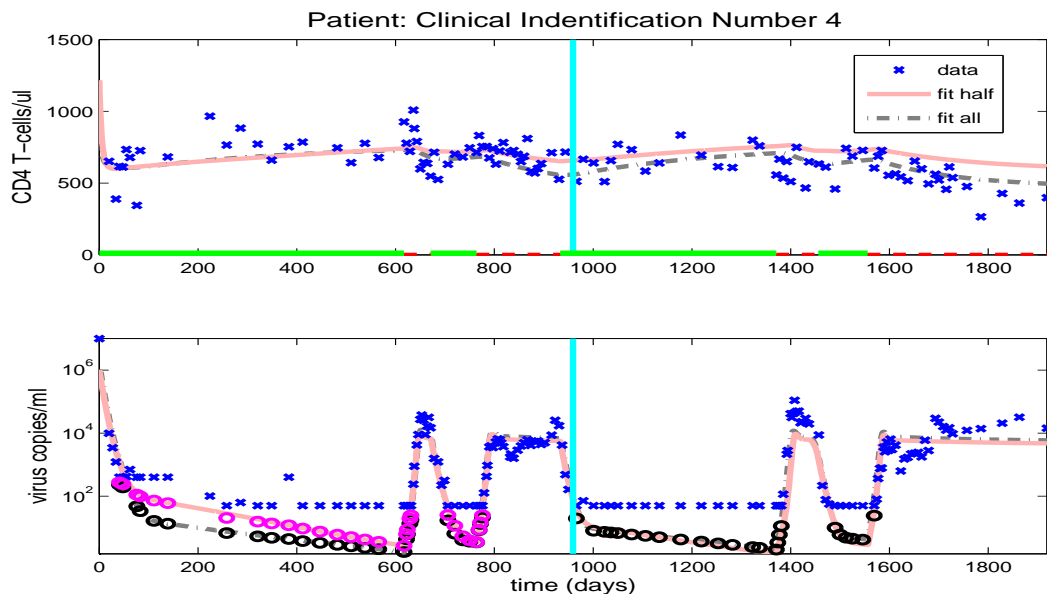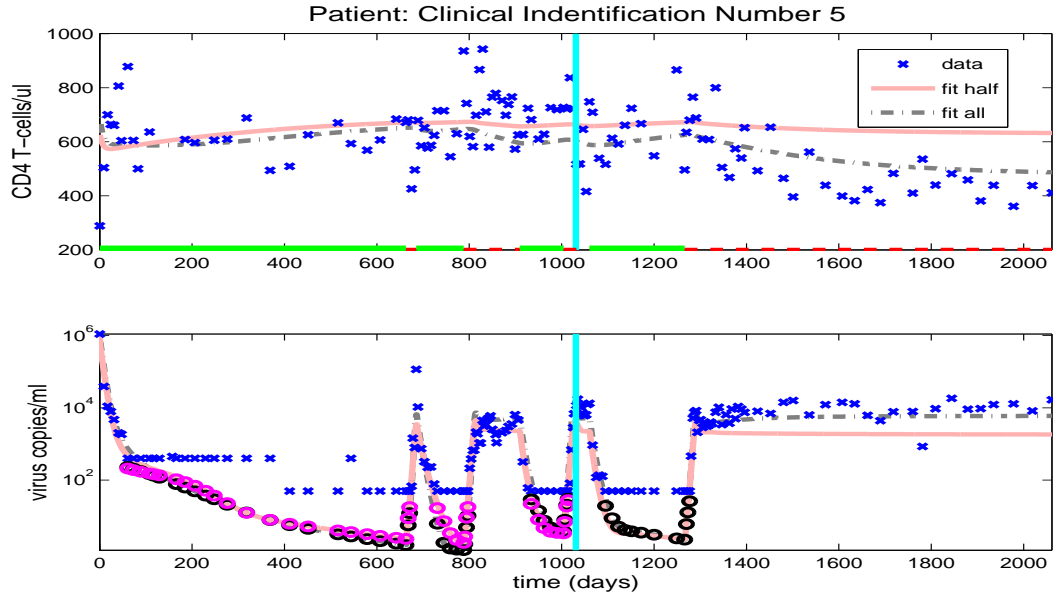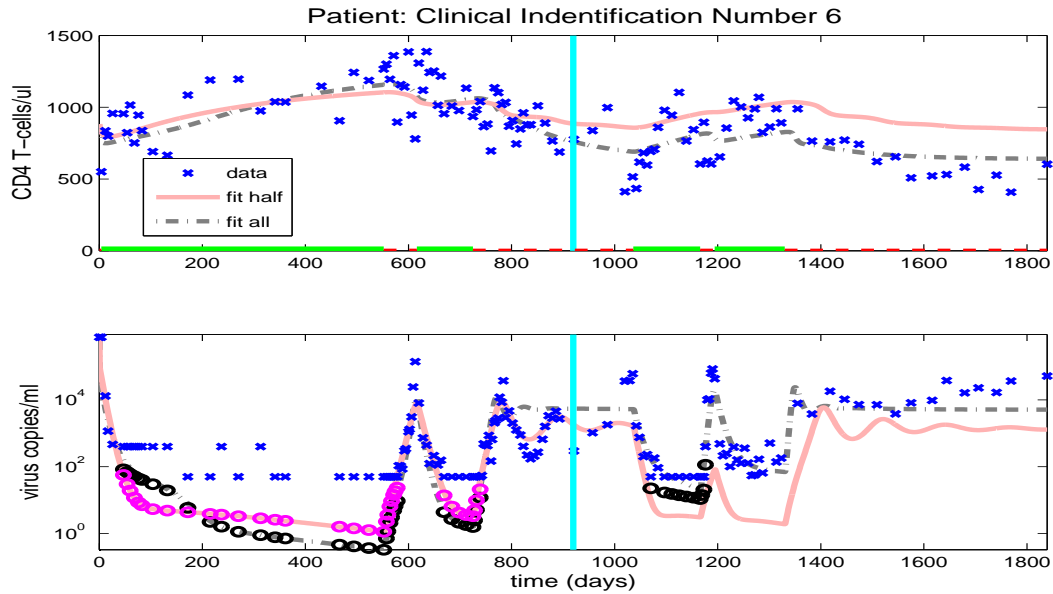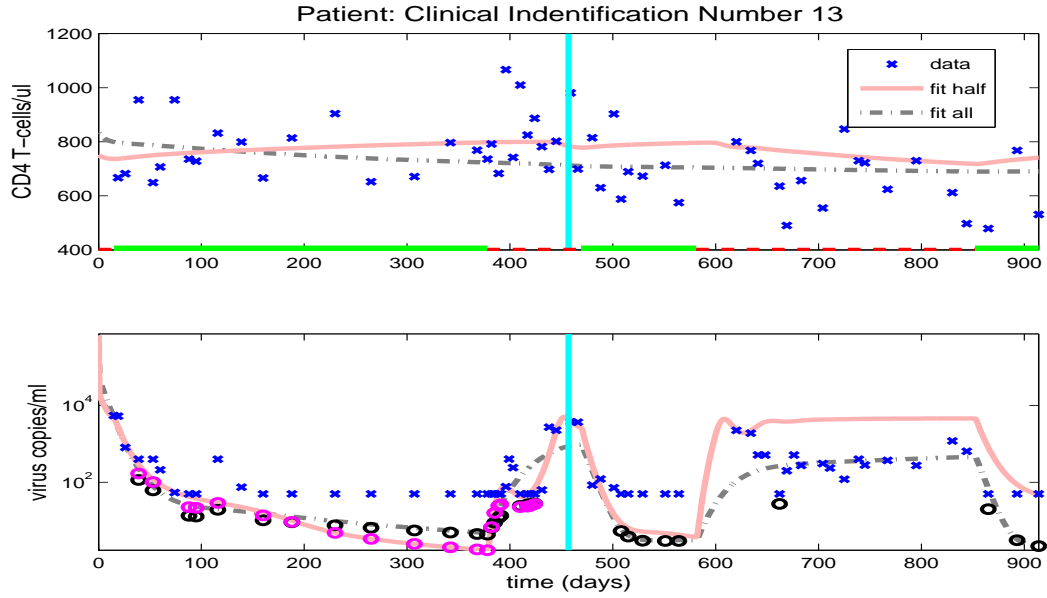

Figure 14: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.
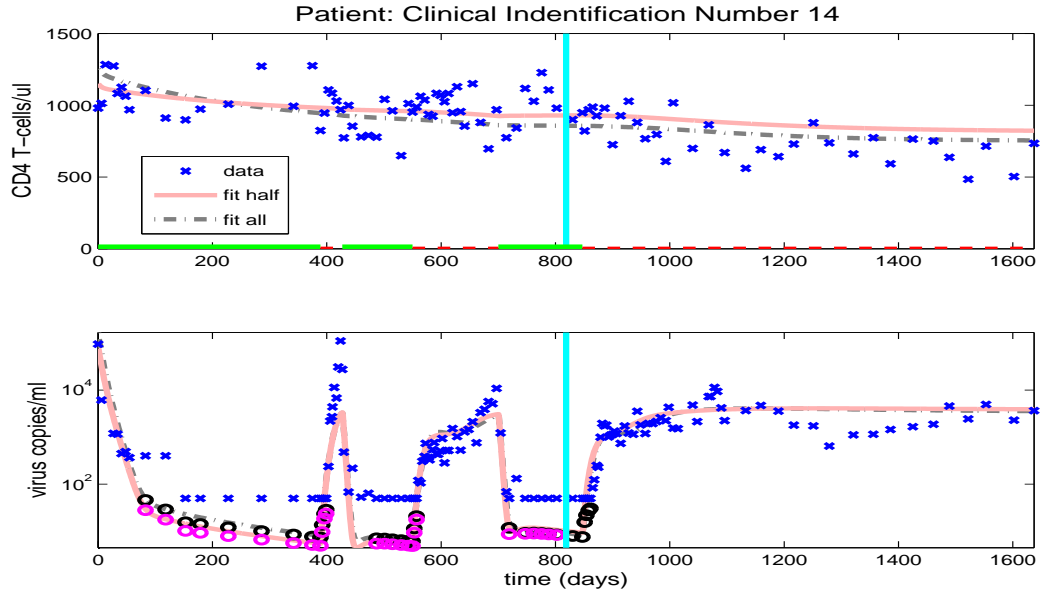
Figure 15: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.
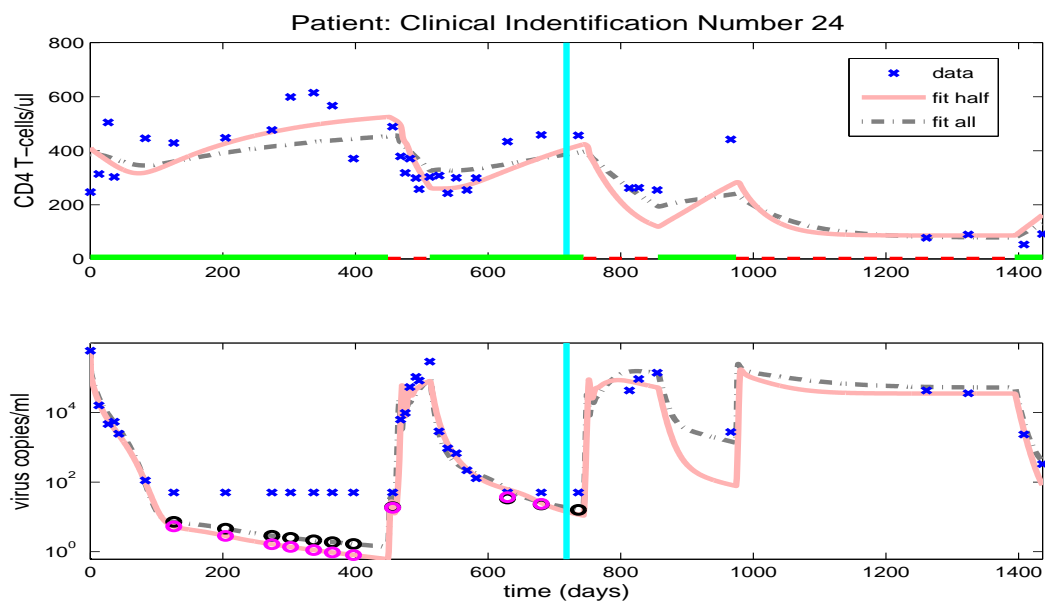

Figure 16: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.

Figure 17: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.
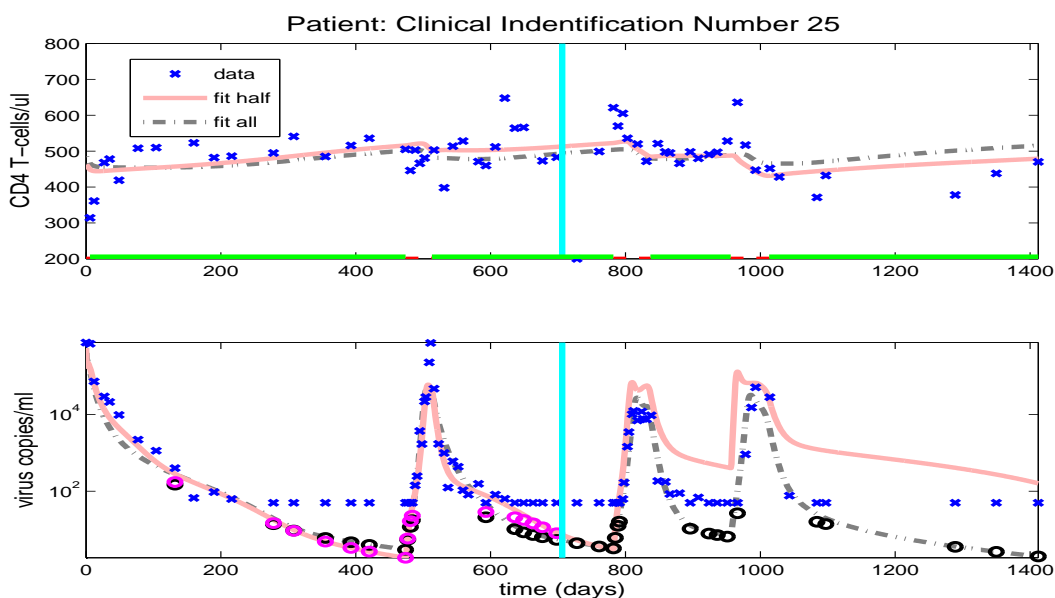

Figure 18: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.
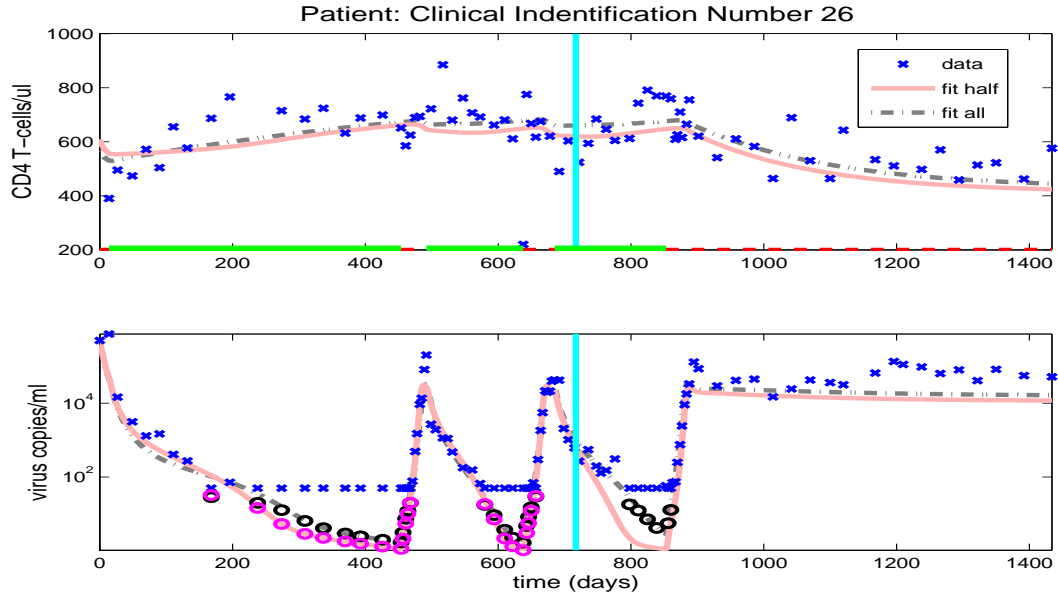
Figure 19: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.

Figure 20: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.
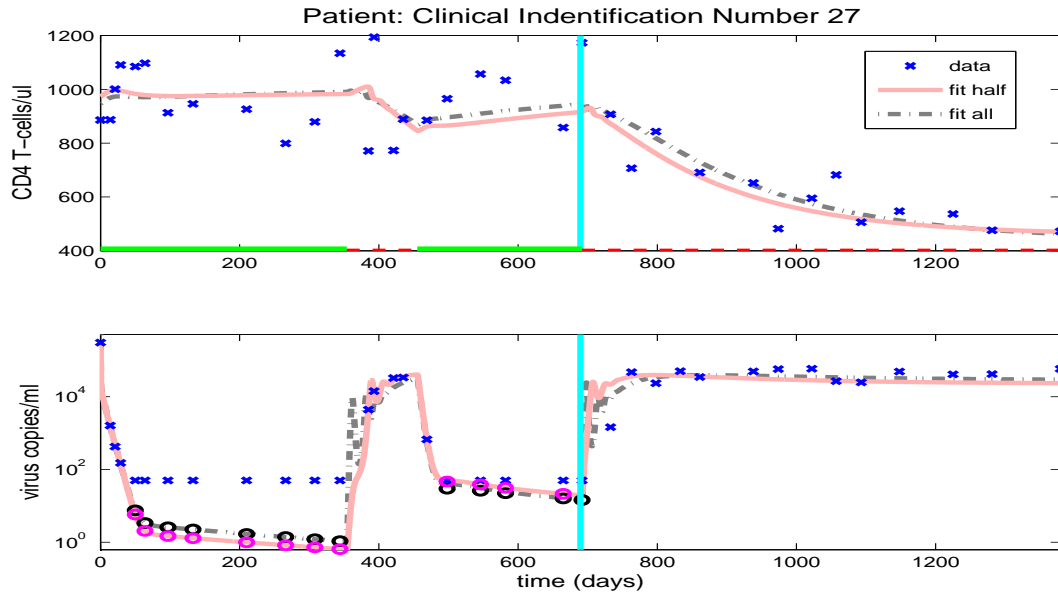
Figure 21: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.

Figure 22: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.
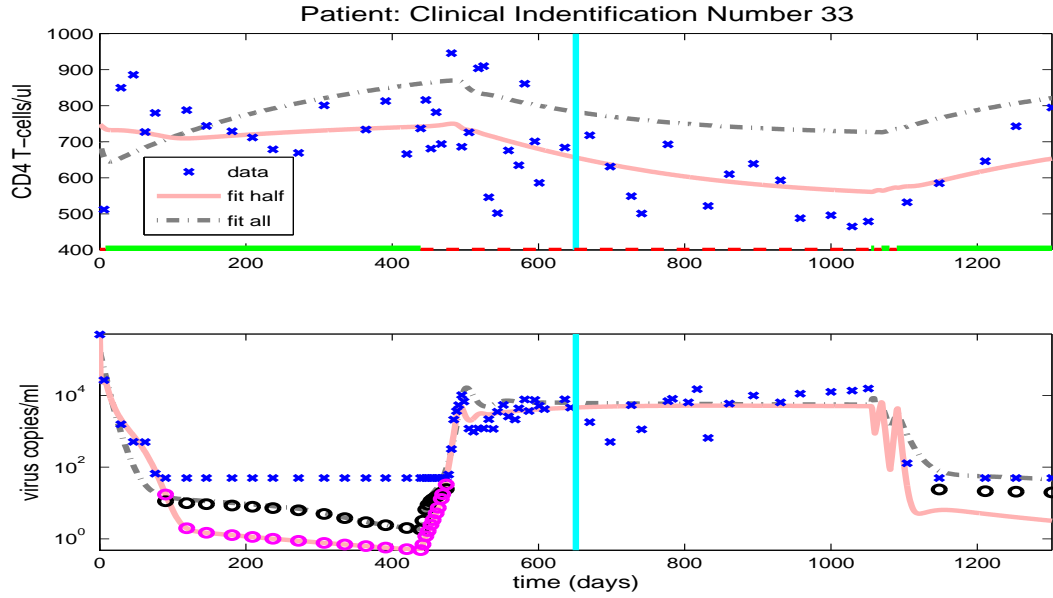
36

Figure 23: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.
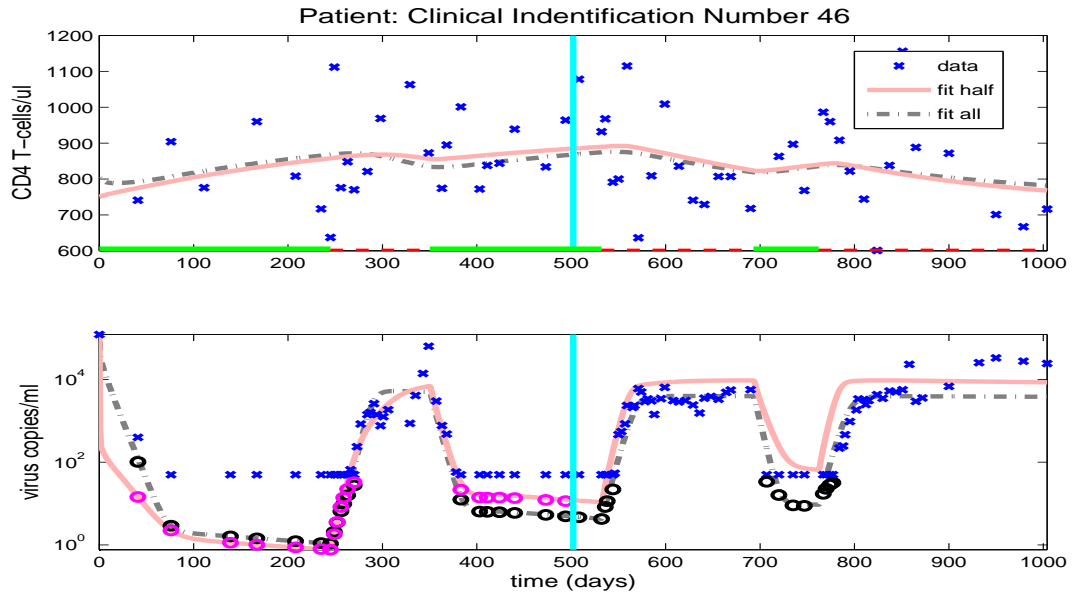

Figure 24: Model fits to data ('x') with parameters estimated from half longitudinal data (solid line) or full data (dash-dot line). Circles denote the predicated values of the censored data and the vertical line delineates between the two halves of the longitudinal data.